REPORT RESUMES

ED 016 230 48 AL 001 027

A GENERATIVE DESCRIPTION OF THE ENGLISH SUBJECT TAGMEMES.

BY- BECKER, ALTON LEWIS
MICHIGAN UNIV., ANN ARBOR,CTR.FOR RES.LANG.AND BEH

REPORT NUMBER BR-6-1784 PUB DATE SEP 67
CONTRACT OEC-3-6-061784-0508
EDRS PRICE MF-$0.75 HC-$7.32 181P.

DESCRIPTORS- TRANSFORMATION GENERATIVE GRAMMAR, *PHRASE STRUCTURE, *SENTENCE STRUCTURE, FORM CLASSES (LANGUAGES), *TAGMEMIC ANALYSIS, ENGLISH, TRANSFORMATIONS (LANGUAGE),

ONE OF FOUR THESES WHICH FORM THE "SUPPLEMENT TO STUDIES IN LANGUAGE AND LANGUAGE BEHAVIOR, PROGRESS REPORT V," SEPTEMBER 1967, THIS STUDY WAS DESIGNED TO RESEARCH (A) WHAT SORT OF INFORMATION ABOUT GRAMMATICAL CONSTITUENTS IS NECESSARY IN A GRAMMAR, AND (B) HOW THIS INFORMATION CAN BE PRESENTED IN A GENERATIVE TAGMEMIC GRAMMAR. THE FIRST CHAPTER DESCRIBES THE NOTION OF "SUBJECT" IN ENGLISH ACCORDING TO SEVERAL PROMINENT LINGUISTS AND PROCEEDS TO DISCUSS FOUR TYPES OF GRAMMATICAL UNITS--(1) GRAMMATICAL FORM (SUBJECT, OBJECT), (2) GRAMMATICAL MEANING (AGENT, GOAL, INSTRUMENT), (3) LEXICAL FORM (NOUN, VERB), AND (4) LEXICAL MEANING (ANIMATE, MALE, HUMAN). CHAPTER TWO PRESENTS R. LONGACRE'S MODEL OF TAGMEMICS WHICH THE AUTHOR CLAIMS IS ADEQUATE FOR DISCERNING SURFACE STRUCTURE (GRAMMATICAL AND LEXICAL FORM), BUT INSUFFICIENT FOR DEEP STRUCTURE ANALYSIS (GRAMMATICAL AND LEXICAL MEANING). THE THIRD CHAPTER INVOLVES A DESCRIPTION OF ENGLISH SUBJECT TAGMEMES AS WELL AS A DETAILED DISCUSSION OF CLAUSE-LEVEL GENERATIVE RULES. THE FINAL CHAPTER CORRELATES THE FOUR TYPES OF GRAMMATICAL UNITS (TAGMEMES) WITH THE DESCRIPTIVE RELEVANCE OF GRAMMATICAL LEVELS (WORD, PHRASE, SENTENCE), AND THE SUFFICIENCY OF THE GRAMMAR IN PROVIDING INFORMATION FOR SEMANTIC INTERPRETATION (DEEP STRUCTURE) IS DISCUSSED. THIS STUDY WAS PUBLISHED AT THE CITY CENTER BUILDING, 220 EAST HURON STREET, ANN ARBOR, MICHIGAN 48108. (FB)

BR 6-1784
PA 48

THE UNIVERSITY OF MICHIGAN

ED016230

Studies in Language and Language Behavior

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

A Generative Description of the English Subject Tagmemes

Alton Lewis Becker

TE000125

CENTER FOR RESEARCH ON LANGUAGE AND LANGUAGE BEHAVIOR

AL 001 027

A Generative Description of the English Subject Tagmemes[1]

Alton Lewis Becker

The problem studied in this work is that of explaining the notion subject in a grammar and describing the realization of this theoretical construct in the grammar of a particular language, English. 2 major questions are examined: (a) What sort of information about grammatical constituents is necessary in a grammar, and (b) how may this information be presented in a generative tagmemic grammar?

In Chapter 1 several descriptions of the notion subject in English are examined, including those of Otto Jespersen, Charles Fries, Eugene Nida, Noam Chomsky, and Charles Fillmore. The relevance of 4 different kinds of grammatical information in the work of each is demonstrated. These 4 kinds of information or aspects of grammatical units are (a) grammatical form (e.g., subject, object, complement, etc.), (b) grammatical meaning (e.g., agent, goal, instrument, etc.), (c) lexical form (e.g., noun, verb, adjective, etc.), and (d) lexical meaning (e.g., animate, human, male, singular, etc.).

In Chapter 2 Robert Longacre's generative model of tagmemics is described and modified. By attempting to write conjoining rules within Longacre's model, several weaknesses are revealed. The major problem is that Longacre's grammar provides information chiefly about grammatical and lexical form (surface structure) and is deficient in describing grammatical and lexical meaning (deep structure). A modified version of Longacre's model is then presented. In this modified model, surface forms are generated within 2-dimensional matrices in which the vectors are categories of grammatical meaning and syntagmeme types.

In Chapter 3 the modified model is applied to the description of English subject tagmemes. Clause level generative rules are discussed and illustrated. Several features of clause level rules are examined in detail: anaphoric deletion, passivity, rank (or focus), lexical subcategorization, dummy tagmemes, and linear vs nesting recursiveness.

In Chapter 4 the descriptive relevance of grammatical levels (e.g., word, phrase, clause, sentence, paragraph, etc.) is demonstrated. The strings generated by clause level rules in Chapter 3 are given phrase level and (briefly) word level exponents. Finally, the sufficiency of the grammar in providing information for the semantic interpretation of grammatical strings is discussed and illustrated.

The study shows the importance of all 4 aspects of grammatical constituents (i.e., tagmemes) and the considerable generative capacity of a tagmemic grammar.

Footnote

[1] This dissertation was supported by the U. S. Department of Health, Education, and Welfare, Office of Education (Contract OEC-3-6-061784-0508), under the provisions of P. L. 83-531, Cooperative Research, and the provisions of Title VI, P. L. 85-864, as amended. This thesis is one of four which have been submitted in a complete form to the Office of Education as *Supplement to Studies in language and language behavior*, Progress Report V, September 1, 1967.

Chapter I

INTRODUCTION TO THE PROBLEM

In this chapter we shall attempt to define the problem to which the following chapters are a partial solution. This problem is that of explaining the notion subject in a grammar and describing the realization of this theoretical construct in a particular language, English. We shall approach this problem by first examining how several grammarians have defined the notion subject in various ways, particularly for English, and then, in Chapter II, we shall introduce an alternative solution. Chapters III and IV include a partial grammar of English in which this alternative solution, based on Robert E. Longacre's generative tagmemic model (Longacre, 1964), is demonstrated in detail.

Before turning to an examination of the work of Jespersen, Fries, Nida, Chomsky, and others, let us first establish the questions we want to ask of each, for merely to list their various observations and analyses would neither reveal their contributions to the grammar presented in later chapters nor be particularly interesting to the reader. In this study of the notion subject, we shall want to know how each defines (explicitly or implicitly) this notion. Furthermore, we shall want to know what terms each uses to define it, i.e., what sorts of linguistic units and constructions made up of these units seem sufficient to each for the description of the notion subject. These questions reflect the assumption that grammatical description is based on the initial activity of identifying and defining in some way units and constructions--in tagmemic terms, tagmemes and syntagmemes; in transformational terms, strings and constituents. It seems clear that the sorts of units and constructions one derives from this initial analysis will in large part determine the sort of grammar one writes; its strengths and weaknesses often rest on the analytical model (intuitive or explicit) employed. To bring in a metaphor from architecture, the material from which a structure is to be built determines in very important ways the sort of structure one can design and build.

1.1 Aspects of Grammatical Units

It seems reasonable to assume that a grammatical unit has at least four basic aspects, the specification of which is necessary to a grammatical description. That is, constituents of a grammatical string carry at least four kinds of information necessary for grammatical analysis. For instance, consider the following sentence:

(a) The little boy walked to the store to get his mother some butter.

It is argued here that a grammatical description of the subject of this sentence will include as a minimum descriptive requirement four sorts of information about that subject. First, a grammatical description will provide the information that the subject is "the little boy", not "his mother" or "some butter" or some other part of the sentence. That is, the string "the little boy" has a grammatical relation to the rest of the sentence (or some part of it) that we can define in some way as subject of the sentence. We presently know of only one major grammarian, Lucien Tesnière, who suggests that the notion subject is not grammatically relevant. He writes:

> Se fondant sur des principes logiques, la grammaire traditionelle s'efforce de retrouver dans la phrase l'opposition logique entre le sujet et le prédicat, le sujet etant ce dont on dit quelque chose, le prédicat ce qu'on en dit. (Tesnière, 1966, p. 103)
>
> En effet tous les arguments qui peuvent être invoqués contre la conception du noeud verbal et en faveur de l'opposition du sujet et du prédicat relèvent de la logique formelle a priori, qui n'a rien à voir en linguistique. (p. 104)

All other grammarians seem to accept the distinction between subject and predicate, (though not necessarily as any more important than that between predicate and object) though Fillmore partially agrees with Tesnière and describes the subject-predicate division as part only of the surface structure of the sentence:

> There are reasons, I have suggested, for questioning the deep-structure validity of the traditional division between subject and predicate, a division which is assumed by some to underlie the basic form of all sentences in all languages. We find ourselves agreeing with Tesnière, who argues that the separation of subject and predicate is an importation into linguistic theory from formal logic which is not supported by the facts of language, and that it is a division which obscures the many structural parallels between 'subjects' and 'objects'. The observation many scholars have made about surface differences between predicative and determinative syntagms do not in any way require us to believe that the 'subject'-'predicate'

> division plays a part in the deep-structure syntactic relations among constituents of sentences. (Fillmore, 1967, pp. 24-5)

As will be pointed out later (2.3), we agree with Fillmore's distinction and will describe the grammatical relation subject as a surface phenomenon.

Secondly, we can identify "the little boy" as the actor or agent in the sentence (a)--the one who "walked". Clearly, the identification of this aspect of the string "the little boy" is not identical to its identification as subject. In none of the following sentences are the subjects also agents:

Walking is good for you.

The store is on Miller Street.

The butter is wrapped in aluminum foil.

That is, merely knowing that X is the subject of the sentence S does not provide sufficient information about X to allow us to determine that X is agent, action, location, goal, or any one of a number of categories of this sort. Grammarians differ widely in their opinions about the relevance of this information; some, including Pike[1] and Fillmore[2], make it central to grammatical description, others, including Chomsky and Longacre, either fail to consider it at all or give it only minor theoretical importance.[3]

Thirdly, we can identify "the little boy" as a noun phrase, i.e., as being an endocentric construction having as its head or nucleus a particular class of lexical forms, nouns. Clearly the information that X is a noun phrase is not derivable from its identification as subject or agent. For example, in the first example above (Walking is good for you.) the subject is neither agent nor noun phrase. Constructions other than noun phrases (e.g., verb phrases, adjective phrases) can manifest both subject and agent in a sentence. For example:

Merely speaking is not enough.

The wicked will perish.

Hence, phrases headed by various lexical form classes (NP, VP, AdjP, etc.) can be identified as units in a construction. As far as I can see, all English grammarians find this information relevant, though I shall suggest below that the specification of this aspect of linguistic units is of minor importance in purely grammatical description, i.e., that it is more a matter of surface than deep structure.

Fourthly, "the little boy" can be identified as having the same referent as "his" later in the sentence. "The little boy" and "his" form an equivalence chain whose domain here is the whole sentence.[4] If the pronoun had been "her" (i.e., The little boy walked to the store to get her mother some butter.), then either there is an error in agreement or the domain of the equivalence chain extends beyond the sentence. Such information can be established formally[5] and is of great importance in discourse analysis and literary analysis. Only a few grammarians (e.g., Pike, Harris) have considered such information relevant to grammar proper. In analyzing the structure of linguistic units larger than clauses, however, the identification of such chains becomes very important, as, for example, in pronominalization and conjoining. (See below, section 2.3.)

These four aspects of grammatical units will be discussed in much more detail as we proceed through the investigation of several previous English grammars; so far they have been hardly more than enumerated. They can be represented in a simple matrix:

	Grammar	Lexicon
Form	A (e.g., Subject)	C (e.g., Noun Phrase)
Meaning	B (e.g., agent)	D (e.g., single male human, etc.)

Diagram 1: Aspects of Grammatical Unit

This display reveals two basic tagmemic assumptions: 1) that all grammatical units are form-meaning composites and 2) that tagmemes are correlatives of syntactic slot and lexical filler. While previously these two assumptions have often been considered identical, they are here considered as separate distinctions, defining the two dimensions of the complex grammatical unit: form: meaning vs. grammar:lexicon.[6]

The question then arises as to whether all four of these aspects of a grammatical unit are necessary: are we assigning too much structure to these units?[7] In the course of this study it is hoped that all four can be shown necessary to grammatical description. At present we ask the reader to suspend judgment until the motivation for each is established below; each proves necessary for describing particular features of grammatical behavior.

1.2 Grammatical Constructions

Let us now turn to the second basic question about the grammars to be examined: what general sorts of constructions does each grammarian identify--i.e., how does each grammarian describe the relations of grammatical units (constituents) to each other? There seem to be three basic ways of describing grammatical constructions, though there are many variants of each. I will label these:

A. String constituent constructions

B. Immediate constituent constructions

C. Center-adjunct (or Nucleus-satellite) constructions

Each of these implies a kind of phrase structure grammar, and as has been adequately demonstrated by Postal (1964) and others, phrase structure grammars are in themselves necessary but not sufficient to an adequate grammatical description. We will accept here most of Postal's conclusions, reserving the right to reject many details of his argument, particularly those concerning tagmemics. However, our immediate purpose here is the examination and classification of several grammars, not the justification of a particular grammar. An identification of different kinds of phrase structure underlying each will help to bring out some of the insights and weaknesses of each.

1.2.1 String Constituent Analysis (S.C.)

Let us refer again to example (a):

(a) The little boy walked to the store to get his mother some butter.

A string constituent grammar (of the sort proposed in Longacre, 1960, and earlier called "Serial expansion" in Pike, 1967, p. 244) divided this sentence into four constituents at clause level:

Subject	Predicate	Direction	Purpose
The little boy/	walked/	to the store/	to get his mother some butter.

This kind of segmentation implies two assumptions about grammatical structure: 1) that the constituents are relevant to clearly defined grammatical levels, and 2) that the constituents are tagmemes (or something resembling tagmemes) derived by possibilities of substitution. That is, strict string constituent analysis isolates functional parts of constructions where substitution from a class of forms (fillers) is possible. These substitution forms themselves may be constructions with a distinctive internal structure representing, usually, a different grammatical level; in the sentence above, the subject slot is filled by a noun phrase, the predicate by a verb phrase, the direction by a prepositional (relator-axis) phrase, the purpose by what we shall call below a nexus. (These are very gross distinctions. A detailed analysis follows in Chapters III and IV.)

Continuing with the string analysis of the example above, we observe that the order of segmentation quite clearly defines these levels:

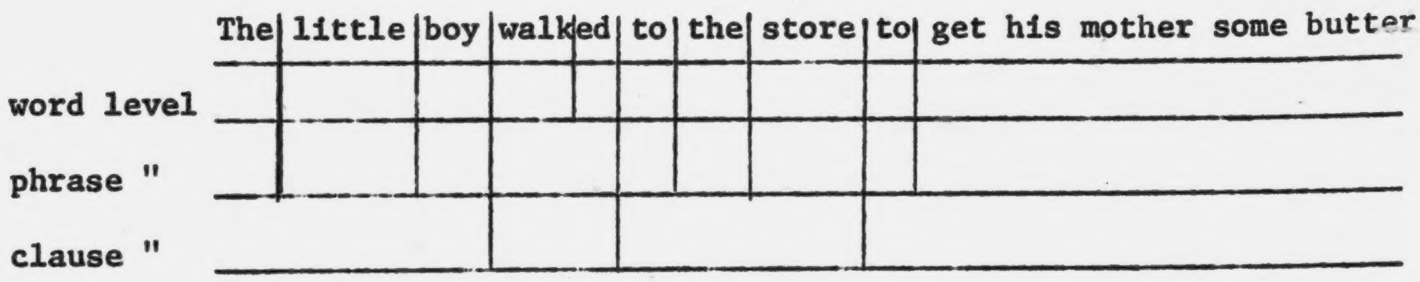

The
little
boy
walk
ed
to
the
store
to
get his mother some butter
word level
phrase "
clause "

Diagram 2: String Constituent Segmentation

The fact that the purpose tagmeme cannot be exhaustively segmented within the dimensions of Diagram 2 is due to the fact that it is headed by a clause level structure itself, illustrating the necessity in a grammar for properties beyond those provided by a phrase structure grammar.

1.2.2 Immediate Constituent Analysis (I.C.)

Immediate constituent analysis assumes a ranking of the various cuts, not in terms of grammatical levels but in terms of orders or degrees of relationship between the constituents. I.C. analysis in practice is strongly binary, separating one constituent from the remainder. The example sentence would be segmented somewhat as follows:

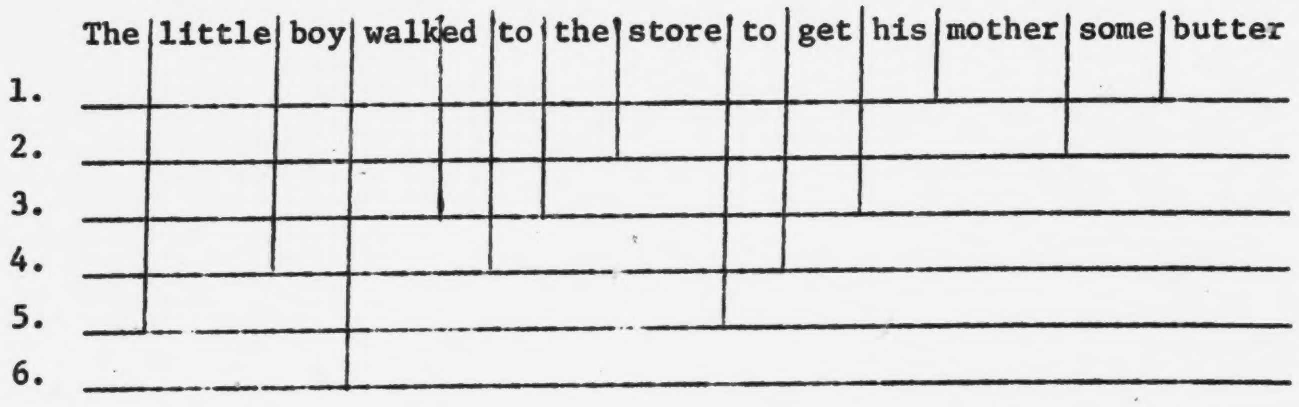

The
little
boy
walk
ed
to
the
store
to
get
his
mother
some
butter
1.
2.
3.
4.
5.
6.

Diagram 3: Immediate Constituent Segmentation

The clause is divided into NP and VP; then the NP is divided into

Determiner and NP (remainder), the VP into final adverbial and VP (remainder), and then these constituents are in turn cut. The levels here (1-6) represent not clearly defined and relevant levels of a grammatical hierarchy but a ranking of the degrees of relationship between constituents. Pittman (1948) has termed this ranking of constituents the "affinity" of I.C.'s and described one constituent as the "pertinent environment" of the other, arguing that in the sentence,

Eat your bread.

the relation of "your" to "bread" is easier to explain than the relation of "your" to "eat".[9] The goal of I.C. analysis is defined by Wells (1947) as follows:

> §18. This is the fundamental aim of I.C. analysis: to analyze each utterance and each constituent into maximally independent sequences--sequences which, consistently preserving the same meaning, fit in the greatest number of environments and belong to focus-classes [i.e. substitution classes] with the greatest possible variety of content.

Hence, as a means (somewhat oversimplified) of distinguishing I.C. analysis and S.C. analysis we may say that I.C. analysis bases segmentation on the need to find grammatically relevant lexical classes (or lexical forms, in Diagram 1) while S.C. analysis bases segmentation on the need to find grammatically functional slots (or grammatical forms, in Diagram 1). Each system does, however, accomodate the central interest of the other, I.C. by describing the distribution of lexical forms in constructions, S.C. by listing fillers (lexical forms) of functional slots. However, because each system of analysis begins with a different criterion for segmenting utterances (each representing a different aspect of the grammatical unit), each comes up with a different grammatical description.

1.2.3 Center-Adjunct Analysis (A.C.)

The third way of describing grammatical constructions--center-adjunct or nucleus-satellite analysis--shares features of both string constituent analysis and immediate constituent analysis. With the former it shares the assumption that analysis be based on the distinction between subordinate (endocentric) relations and interordinate (exocentric) relations between grammatical units. That is, it insists on a major distinction between such modifier-head relations as

The barking dogs

and interordinate relations like the subject-predicate relation, as in

The dogs bark.

The difference between these two kinds of relationships is not revealed by a strict I.C. analysis.

On the other hand, center-adjunct analysis shares with I.C. analysis an emphasis on specifying the ranking or degrees of relationship between constituents, what Pittman called the "affinity" of constituents. As we shall see below in the example, C.A. analysis peels off layers of constituents in such a way as to retain the elementary parts of the sentence, Subject-Predicate, or Subject-Predicate-Object. (Note: Harris defines S so that all S's are S-P-O where O may be Ø.)

The formally most rigorous example of C.A. analysis is to be found in Harris's String Analysis of Sentence Structure. Though Harris calls his model "string analysis", it is clearly not the kind of string analysis proposed by Longacre[10] and hence is differentiated here as C.A. (center-adjunct) analysis. The major difference between the two models is that S.C. analysis isolates relevant "vertical" layers or levels of structure (i.e., word level, phrase level, clause levels, sentence level, etc.) while C.A. analysis does not.

Harris (1962) describes the procedure for C.A. analysis as follows:

> We now consider each sentence S as a sequence of morphological word-categories (or sub-categories, or disjunctions of categories, or rarely sequences of categories)s_i. When we are given an arbitrary sentence S, we isolate out of it the elementary part S_o by asking what contiguous sequences of the s_i can be exised, one sequence at a time, by operations of general or nearly general applicability, the residue of S after each excision being still a sentence of the language. (p.22)

For example, using (a), we get:

S = The little boy walked to the store to get his mother some butter.

Excise:	Leaving a sentence:
a. to get his mother some butter	The little boy walked to the store.
b. to the store	The little boy walked
c. little	The boy walked = S_o

The order of excising is generally irrelevant, except that the remainder must be a sentence (S_n) with the same properties of occurrence as the original sentence (S).[11] The excised constituents (Z_n)

are then themselves reduced to elementary parts ($Z_{n/o}$) in the same way, i.e., "by excising successively various word sequences included in Z_i, the residue of Z_i after each excision having the same properties of occurrence that Z_i had." (p.23) Only (a) in the example above can be reduced further:

Z_i = to get his mother some butter

Excise:	Leaving the residue (also a Z):
d. his mother	to get some butter
e. some	to get butter = $Z_{i/o}$

Like S.C. analysis, and unlike I.C. analysis, A.C. analysis defines an elementary sentence within the original sentence, allowing the analyst to relate all sentences having the same elementary structure (S_o).[12] It also brings to the foreground the notion of endocentricity or modifier-head. In Tagmemic grammars (based, in general, on a complex version of the string constituent model) these endocentric relations are marked by the signs + and ± (and certain elaborate variations of these signs involving brackets and ties between signs) before constituents; in Transformational grammars they are marked by parens and optional transforms. Such relations do not appear in strict I.C. analysis; that is, in a tree diagram. What Harris has done for each S is excise all constituents which in a tagmemic grammar would be dominated by the sign ±.

1.2.4 Other Construction Types

The three methods of analysis are, of course, closely related to the various aspects of grammatical units discussed above:

	Grammar	Lexicon
Form	A (e.g., Subject)	C (e.g., Noun Phrase)
Meaning	B (e.g., agent)	D (e.g., single, male human, etc.)

Diagram 1: Aspects of Grammatical Unit

I.C. analysis emphasizes C (lexical form) although bringing in vocabulary of A at times (e.g., subject, predicate, object). S.C. analysis (in its early form) and C.A. analysis emphasize both A (grammatical form) and C (lexical form), i.e., both functional slot and substitution class. It is intriguing to ask now whether analytical procedures based on B or D (grammatical or lexical meaning) are possible.

Clearly aspect B (grammatical meaning) can be and has been used as a means of analyzing grammatical strings. The notion is present in Jespersen and central in the recent work of Fillmore (1967), what he calls "case" grammar. It has been demonstrated in a tagmemic grammar in Pike (1964a and 1967, p. 576-77, 607) and Becker (1967). In this approach, constituents are marked by such labels as agent, goal, motion, location, etc. The sentence,

> The little boy walked to the store to get his mother some butter.

might be derived from the string (or proposition)

> Agent Motion Direction Purpose.

Though there is as yet no formal and explicit procedure for analyzing a sentence into these constituents, there is a heuristic based on question transformations or the substitution of certain category labels for each constituent. That is, choice of question word often marks "cases" or categories of grammatical meaning:

a. Who walked to the store to get his mother some butter?
b. How did the little boy go to the store to get his mother some butter?
c. Where did the little boy walk to get his mother some butter?
d. Why did the little boy walk to the store?

Here (a) evokes as a substitutable constituent actor, (b) either motion or manner, (c) location or direction, and (d) either purpose or cause. Expanded forms of these questions can evoke answers less ambiguously; notice that each of the following questions puts the sentence into the form,

Question-Substitute + Be + Category Label + S

a. Who was the person who walked to the store to get his mother some butter?
b. What was the means by which the little boy went to the store, etc.?
c. What was the place to which the little boy walked, etc....?
d. What was the purpose for which the little boy walked to the store?

Formalizing the rules for this procedure of analysis is no more difficult than formalizing the rules for question transformations or relative clause transformations: both demand that information like the following be available:

(a) "boy" may be replaced by "who" or "person" or "he"
(b) "purpose for which" is equivalent to "why"
(c) "place to which" is equivalent to "where"

It will be one of the major goals of this thesis to show the relevance of grammatical equivalence and to provide a much more explicit definition of "grammatical equivalence" based on the specification of aspect B (and to a lesser extent aspect D) of grammatical units.

Now, however, let us first examine several English grammars in relation to the three theoretical questions raised here:[13]

1. Which of the four aspects of grammatical units are used in each particular grammar? Which seem central to the analysis?
2. What kind(s) of constructions (relations of units) are defined?
3. How is the notion "subject of a sentence" defined (explicitly or implicitly)?

1.3 Previous Descriptions of the English Subject

1.3.1 Otto Jespersen

> "In Syntax Meaning is Everything"
>
> -*A Modern English Grammar*, Vol. IV, p.91

It is difficult to begin with Jespersen both because his work is extremely detailed and, in many ways, far more complete than others we shall examine and also because his manner of presentation is almost entirely taxonomic and unmotivated formally. Jespersen is very modern in his analysis, for instance, of "the barking dog" and "the dog barks" as identical in what we would call now "deep" structure. At the same time, because his grammar was not at all *generative*, there is no way of evaluating such of his notions as that subject and object in a sentence (nouns) are primaries, while the predicate is a secondary. Such distinctions are often left unsupported: what descriptive advantages accrue from them? Why is the distinction necessary?

Of the four aspects of grammatical units discussed above, Jespersen touches on all of them in some way. Aspect A (grammatical form) is described in detail. Jespersen recognizes such functions as subject, predicate, and object and defines them partly in terms of form:

> "The subject is the primary most intimately connected with the predicate in the form which it actually has in the sentence with which we are concerned." (1927, p.206)

In order to identify the subject in a given sentence, Jespersen transforms the sentence to who (or what) followed by the verb in the form used in the sentence. Thus, the sentence,

The boy walked to the store.

is transformed to

Who walked?

The answer, "The boy", is the subject. I think it would be said today that this process is not entirely formal, that writing the rule for the question transformation requires information not provided in Jespersen's simple rule. E.g.,

The boy walked Mrs. Brown's dogs in the park.

Here we have to specify for the question transformation whether who or what is to be substituted:

Who walked? The boy.

What walked? The dogs.

In essence, Jespersen's rule for finding subject specifies the substitution of forms unmarked for case by forms marked for case. That is, the boy carried no case marker, while substitute forms (who, he) do. And in order to make these substitutions properly we must draw upon information about sentences which is not formally marked in the surface structure: e.g., that the form to be substituted is human and male.

As Jespersen noted, (1927, p.228) his rule for finding subject would not work in all instances:

He happened to fall.
He was sure to fall.
He was believed to fall.

These transform to the questions,

Who happened?
Who was sure?
Who was believed?

The "answers" to these "questions" are not clear. Jespersen argues (1927, p.228) that "there cannot be the slightest doubt as to the grammatical subject: it is *he*." He goes on to argue that *he* is not just the subject of the predicates (*happened*, *was sure*, *was believed*) but of the whole sentence. In a *notional* sense, *he* is the subject of the "predicate" *to fall* and the notional subject of *happened* is the nexus *he-to-fall*. Hence, there seem to be something like surface subjects and deep subjects in these sentences. Unfortunately, Jespersen did not do more than point out this interesting difference--i.e., he did not go on to discuss or formulate its implications for a theory of language which involved a contrast between deep and surface relationships. Here, as in many instances, Jespersen has asked an important question which has held the attention of succeeding generations of lingusits.

Jespersen also noted the difference between grammatical form and grammatical meaning, a difference implied in the previous paragraph, where the formal subject is he and the notional (or logical) subject is the "idea" or nexus he-to-fall, in the sentence,

He happened to fall.

As Jespersen stated, (1933, p.107) "the subject cannot be defined by means of such words as active or agent." He recognized that the subject of a sentence may (in his term) "denote" other things besides agent, e.g., in a passive sentence or in such pairs of sentences as,

The garden swarms with bees./Bees swarm in the garden.

I was in sight of the shore./The shore was in sight.

The "notional" term agent is not equivalent to the grammatical term subject. Furthermore, in sentences where the predicate is a form of be, the subject can usually be distinguished from the predicate nominative because the subject is more definite than the predicate nominative:

Tom is a scoundrel.

A scoundrel is Tom.

In both cases, Jespersen argues, (1934, p.150 ff.) Tom is the subject, the second instance being an inversion. Here our recognition of the subject depends probably on the determiner and that Tom is a proper noun. In all these cases, the specification of subject involves more than just a formal relation with the verb. It involves what has been called grammatical meaning above. In Jespersen's terms, the grammatical subject and notional (or logical) subject of a sentence may not be the same. A structural description of sentences like,

He happened to fall.

He is easy to find.

requires that both these aspects of grammatical units be considered. In both sentences He is the grammatical subject, but a more than superficial description of the sentences requires us to show--in some way--that He is also the agent of fall and the goal of find in the deep structure of the sentences.

The aspect of grammatical units which we have labelled lexical form (C) is traditionally known as "parts of speech". Jespersen (1927, p.206) distinguished this aspect of the subject by listing forms which may be "subject of a verb":

a substantive: the man rises

an adjective: the dead arise

a pronoun: he rises

a clause: that he has come is surprising

an infinitive: to hear him is a great pleasure.

Using the tagmemic distinction, these can be seen as the filler-classes of the grammatical function, subject. That these classes are relevant beyond the specification of aspects A (grammatical form) and B (grammatical meaning) is evident in Jespersen's examples:

The man rises.

The dead arise.

Both the man and the dead are subjects (identified by Jespersen's substitution rule for finding subject: Who rises?, Who arise?) and both are agents or actors. However, the man is clearly singular and the dead, plural. (He rises; they arise.) In fact, subjects as agents filled by adjectives are nearly always plural in English. This is a property of the filler class, derivable neither from grammatical form nor meaning.

Furthermore, members of lexical form classes frequently undergo distinct derivational processes (and take on different forms) in different syntactic relations. Each has a distinct set of inflectional and derivational modifications:

act: action, active, actor, etc.
die: death, dead, dier, etc.

Though English words shift quite easily from one form class to another, so that we can freely form verbs from substantives and substantives from clauses (e.g., baby-sitting from baby-sitter from She sits with the baby), there is probably no grammar that can dispose of the notion of form classes entirely. That procrastinate and become are verbs and boy and automobile are nouns is apparently primitive, non-redundant information in a grammar.

That aspect of a grammatical unit which we have called (D) lexical meaning indicates (in part) a particular reference in a discourse, and, as Jespersen pointed out (1923, p.123 ff.), this reference may shift in the course of a discourse. Jespersen describes a number of particular instances of this involving words like father, mother, enemy, home, one, I, you, etc. He called these words "shifters".

The relevance of this insight by Jespersen for the formulation of grammatical rules has been pointed out by Jacobson (1957). It is clear in English, for instance, that _I_ cannot be conjoined to itself in a sentence:

I and I are here.

and when _this_ is so conjoined,

This and this are important.

there are _two different references_ for _this_. Likewise, if there is only one reference for _John_ in the sentence,

John shot John in the foot.

then an automatic reflexive pronominalization must take place:

John shot himself in the foot.

If such phenomena are to be described in a grammar, then this fourth aspect of grammatical units must somehow be specified. In recent years, Jespersen's _shifters_ have become a major focus of attention for linguists--another illustration of his providing the right questions, if not the answers. The problems raised earlier about his rule for finding subject by substitution of pro-forms and his recognition of _he_ as both subject of _be_ and object of _find_ in

He is easy to find.

both seem to center on this problem of specifying reference. That is, there are deep rules in English prohibiting the repetition of a particular lexical form with unchanging reference.

Jespersen identifies three types of grammatical constructions.

The terms associated with these types are rank, nexus, and junction. Jespersen's notion of rank seems to have been motivated by the same goals that Wells (1947) mentioned for I.C. analysis, that one constituent is the "pertinent environment" of another. Like Wells he related his particular ranks (primary, secondary, tertiary) to substitution classes (i.e., lexical forms). Thus, in the phrase,

terribly cold weather,

Jespersen writes (1933, p.78) "weather is determined or defined by cold and cold in its turn similarly determined or defined by terribly. There are thus three ranks: weather is Primary; cold, Secondary; and terribly, Tertiary in this combination." He adds further that "substantives, adjectives, and adverbs habitually stand in this relation to one another", although "in some combination a substantive may be secondary or tertiary, an adjective may be a primary, etc." He goes on to illustrate the use of form-classes (substantives, adjectives, verbs, adverbs, etc.) as primaries, secondaries and tertiaries.

This notion is clearly that which motivates decisions about segmentation in I.C. analysis. Jespersen (1933, p.90) recognized layers (though not levels) of ranking:

> While the distinction between substantives, adjectives, etc., (word-classes, parts of speech) concerns single words in themselves, word-groups as such may be employed in the same ways as single words in different "ranks," and then have to be termed primaries, etc., according to circumstances. The rank of the group is one thing, the rank within the group another. Thus, the group Sunday afternoon, which contains a secondary Sunday and a primary afternoon, may as a whole be
>
> a primary: Sunday afternoon was fine,

a secondary: a Sunday afternoon concert, or
a tertiary: he slept all Sunday afternoon.

His criteria result in nearly identical layers as those identified previously in the I.C. analysis of the sentence (except that Jespersen generally left out articles):

The little boy walked to the store to get his mother some butter

layers:

1) "The little boy" is primary, "walked" and its adjuncts, secondary.

2) "boy" is primary, "little", secondary.

3) "walked to the store" is primary, "to get his mother some butter", secondary, etc.

Jespersen then distinguished two ways that a secondary could be joined to a primary by junction and nexus. These two relations are nearly identical to the endocentric and exocentric relations which, as we have seen above, underlie center-adjunct analysis. The relation of subject, predicate, and object is a nexus, in which subject and object (if there is one) are primaries, predicate secondary. This classification of the predicate as a secondary seems to have been motivated by Jespersen's analogy between pairs like

The barking dog

and

The dog barks.

In the former, "barking" is clearly secondary and, by analogy, may also be in the latter. This is the only apparent reason in the context of Jespersen's works for seeing the predicate as secondary in a sentence. In section 3.2.3 below, however, some further evidence supporting this interpretation will be given.

This fusion of, roughly, I.C. analysis and center-adjunct analysis provided Jespersen with a technique which recognizes two overlapping but distinct syntactic features, the layering of constituents (pertinent environments) and functional relations of major parts of a sentence (subject, predicate, object, etc.). In showing the underlying similarities between junction and nexus (e.g., as in "the barking dog" and "the dog barks") Jespersen suggests a perspective from which the difference between endocentric and exocentric constructions begins to disappear. In section 3.2.3, this suggestion will be developed further.

1.3.2 Charles C. Fries

"The grammar of a language consists of devices that signal structural meanings."

-*The Structure of English*, 1952, p.56

In spite of surface differences and Fries's insistence (1952, p.57) that he, unlike traditional grammarians, "starts from a description of the formal devices that are present and the patterns that make them significant and *arrives at the structural meanings as a result of the analysis*," the resultant grammar is very like Jespersen's. Like Jespersen, Fries includes all four aspects of linguistic units in his description and, also like Jespersen, he conceives of constructions both as center-adjunct patterns and as layerings of I.C.'s. Fries's contribution lies, probably, in his attempt (only partially successful) to formalize the procedures for identifying units and segmenting constructions.

Fries's definition of subject (1952, p.183) brings in all the four aspects we have discussed above:

> The subject itself is simply that Class 1 word that is bound to a Class 2 word to form the basic arrangement of the sentence, and is identified and distinguished from other Class 1 words not by meaning but by certain contrastive arrangements.

This definition can be explained in terms of the four aspects. Grammatical form (A) is expressed in the notion "bound to" in the definition above; elsewhere this notion is called a tie between word classes and is represented by a double-headed arrow: Class 1 <----> Class 2. If we substitute (though the terms don't entirely overlap) Jespersen's term "primary" (in a nexus) for Fries's Class 1 word, Jespersen's intimately connected with for Fries's bound to (see Jespersen's definition above, p.17), their definitions seem nearly equivalent. For Fries, "bound to" meant proportionate in person and number, formal concordance--a proportion revealed for Jespersen by substitution of pro-forms (personal and relative pronouns) which agreed in person and number (and case) with the verb. For both, the subject is discovered via a quasi-transformation:

Jespersen: The boy kicks the dog ---> who kicks? The boy/he.

Fries: The boy kicks the dog --> The boys kick the dog.

(The Class 1 word that covaries with the verb is the subject.)

Though in his earlier work, American English Grammar, Fries (1940, p.140 ff.) like Jespersen makes a distinction between logical and actual subject in such sentences as, It is right to send him.

Here *it* is seen as the actual subject and *to send him* as the logical subject. He does not preserve this distinction in his later work. He does, however, show (1952, p.175) that the relational term subject does not correspond to "the actual facts of a situation in the real world". He thus makes a distinction parallel to Jespersen's between grammatical subject and logical (or notional) subject, when he demonstrates (1952, p.175 ff.) the independence of the relation "subject" from terms like performer of action, receiver of action, the act, time, etc:

> Let us assume, for example, a situation in which are involved a man, a boy, some money, an act of giving, the man the giver, the boy the receiver, the time of the transaction yesterday. With a real situation composed of these factors and a statement to be made containing them all, it is impossible to predict which elements would be grasped in the linguistic expression as "subject" or "indirect object" or "direct object." Any one of the units man, boy, money, give, yesterday could appear in the linguistic structure as "subject".
>
> The man gave the boy the money yesterday.
> The boy was given the money by the man yesterday.
> The money was given the boy by the man yesterday.
> The giving of the money to the boy by the man
>   occurred yesterday.
> Yesterday was the time of the giving of the money
>   to the boy by the man.

He goes on, however, to show (1952, p.177) how the notion, "performer of an act" can be expressed in a number of ways:

> The *men* *built* that tool house very slowly.
> The tool house was *built* *by* *the* *men* very slowly.
> *Their* *building* of the tool house was very slow.

Here is a recognition, much like Jespersen's, of the underlying identity of different constructions, the identity of "the men built" and "their building"--in Jespersen's terms, a nexus and a junction. In a very important step, often overlooked in the era when "meaning" was a phobia for many linguists, Fries shows (1952, p.177) how subjects can "represent at least five different meanings", and how these meanings can be formally determined. The five meanings are:

1. Subject as performer: The dean approved all our recommendations.
2. Subject as that which is identified: One difficulty is the size of the trees.
3. Subject as that which is described: The abstract is very bulky.
4. Subject as that which undergoes the action: The requisition was sent over a week ago.
5. Subject as that "to or for" which the action is performed: Mr. W--- was given the complete file on ______.

Fries (1952, pp.178-80) gives formal rules for identifying these five meanings:

> (a) The subject signals "performer" whenever the Class 2 word [verb] with which it is bound is not one of a special list, or one of the forms of be or get as function words with the so-called "past participles."
>
> (b) The subject signals "that which is identified" whenever the Class 2 word [verb] with which it is bound is one of a special narrow (2b) list (the forms of be most frequently) and this Class 2 word [verb] is followed by a Class 1 word [noun] having the same referent as the Class 1 word [noun] which is "subject."
>
> (c) The subject signals "that which is described" whenever the Class 2 word [verb] with which it is bound is one of the special (2b) list (the forms of be most frequently) and this Class 2 word [verb] is followed by a Class 3 word [adjective].
>
> (d) The subject signals either "that which undergoes the action" or "that to or for which the action is performed" whenever the Class 2 word [verb] to which the subject is bound is the function word be (in its various forms) or get, with the so-called past participle.

A major difficulty with Fries's operational definition of grammatical meaning (our aspect B) is that he didn't go far enough in describing the relation of the categories of grammatical meaning he distinguishes to various kinds of grammatical constructions. For instance, can we write a grammar in which the relation between performer, say, in:

> The men built that tool house very slowly.

and performer in

> Their slow building of that tool house

is explained? Furthermore, what other categories of grammatical meaning underlie a given construction?

In the sentence,

> Yesterday was the time of the giving of the money to the boy by the man.

the word yesterday manifests both subject meaning (b) (that which is identified) and what Fries elsewhere (1952, p.186) calls an "adverbial object" with the meaning time. Is the time meaning lost in the subject relation, while the performer meaning is retained, for example, in

> The tool house was built by the men very slowly.

There seems to be an inconsistency here, an obscuring of the underlying equivalence between the two sentences

> Yesterday was the time of the giving of the money to the boy by the man.
>
> The man gave the boy the money yesterday.

Fries's work, however, remains a pioneering effort to explain and formalize grammatical meaning.

Clearly the third aspect of linguistic units, lexical form (C), is central in Fries's analysis. Syntactic relations and grammatical meaning are both defined in terms of word classes. These classes are derived by substitution in three test frames:

Frame A: The concert was good (always).

Frame B: The clerk remembered the tax (suddenly).

Frame C: The team went there.

Words substitutable for concert in Frame A (and also for clerk and tax in B and team in C) were labeled Class 1 words, words substitutable for was, remembered, and went are Class 2 words, etc. Four classes were established this way, corresponding to the traditional classes, noun, verb, adjective, adverb. Similarly, 15 groups of function words were established--words like a, my, two (Group A, substitutable for the in the frames), those like can, will, has to (Group B, substitutable for may in "The concert may be good"), those like quite, any, too (Group D, substitutable for very in "The concert may not be very good then."), etc. These groups include the traditional closed classes like auxiliaries, determiners, negative, conjunctions, prepositions, relative and interrogative pronouns, and others. These groups differ from the four classes in a number of ways (1952, p.106 ff.):

1) Groups are much smaller sets than classes

2) Minimum single free utterances need include only classes; groups appear most frequently in expanded single free utterances

3) The lexical meaning of words in the classes is "clearly separable from the structural meanings of the arrangements in which these appear." Not so for groups.

4) "In order to respond to certain structural word signals one must know [function words (members of groups)] as items." Words from the four classes often may be replaced by nonsense forms without losing structural meaning; not so for function words.

Once established, these classes and groups are used as the variables in establishing definitions for structural meaning and rules for specifying syntactic patterns.[14] Definitions of structural meaning, as we have seen, take the form:

a) The subject signals "performer" whenever the Class 2 word with which it is found is not one of a special list [e.g., be, become, seem, etc.] or one of the forms of be or get as function words with the so-called "past-participles."

b) The subject signals "that which is identified" whenever the Class 2 word with which it is found is one of a special narrow (2b) list (the forms of be most frequently) and this Class 2 word is followed by a Class 1 word having the same referent as the Class 1 word which is "subject."

Rules for specifying syntactic patterns take the form:[15]

$$\begin{array}{cccccccccc} D & \underline{1^a} & \underline{1^b} & \underline{2\text{-}d} & D & \quad & 1^c & f & D & 1^d \\ & \text{it} & \text{he} & \pm & & & + & E & & + \\ & & & & & & \text{it} & & & \text{it} \end{array}$$

Example:

The library assistant brought the papers and the grades.

The symbols may be interpreted as follows:

D = Any determiner

1 = Any Class 1 word

1^a = The letter exponents indicate whether referents of two Class 1 words are "the same" or "different". Words with the same exponent have the same referent.

$\underline{1^a}$ = The symbols under the figures represent number: $\underline{\quad}$ = singular, + = plural, ± = form that could be singular or plural.

$\underline{1^a}$ = The words under the number symbol indicates the "substitute
it group" to which it belongs (it, he, he/it).

2-d = The symbol following the number representing a Class 2 word indicates the form of that word: -d is preterit or past participle.

f = Any Function word

f_E = Function word from Group E (conjunctions)

For Fries (1952, p.201), the name "subject" is the name of a structural slot in which Class 1 words appear. The structures are established by formal contrast, not by meaning, and "are the signals by which we receive and convey such meanings as 'performer', 'identification', 'that which undergoes the action', 'that for whom an action is performed', and 'that which results from the action'."

However, two interesting kinds of "meaning"--both included in our aspect D (lexical meaning)--remain in his syntactic rules. The letter exponents of Class 1 symbols (e.g., 1^a) indicate whether referents of two Class 1 words are "the same" or "different". And words under Class 1 symbols ($\underset{it}{\underline{1\ }}$) indicate what substitute group the word belongs to. Both of these kinds of information are covert in English and cannot be established from any overt or surface markers in the physical representation of the sentence. We have already seen that information about reference is essential in describing such grammatical processes as conjoining and pronominalization. Likewise, what Fries calls the "substitute group" of a word may also be necessary in writing conjoining rules (in Fries's system); for example, in the sentences:

John built the house with his father.

John and his father built the house.

John and *his father* are both represented in Fries's notation as $\underset{\text{he}}{(1)}$ and hence are conjoinable. While, in the sentences:

John built the house with his axe.

*John and his axe built the house.

John is $(\overset{1}{\text{he}})$ and *his axe* is $(\overset{1}{\text{it}})$ and hence are not conjoinable here.

While agreeing, then, with Fries that information about reference and substitution class are important in a descriptive grammar, we would argue that both are "meaning" (what has been called lexical meaning) and are not established by formal contrast between structures. There are, to adapt Pike's phrase, lexical prerequisites to syntactic analysis.[16]

It is interesting that Fries adds I.C. Analysis to what is *primarily a variation* of center-adjunct segmentation. Unlike Jespersen, who merely suggests the relation of ranking (I.C. analysis) to nexus (center-adjunct analysis), Fries (1952, p.267) specifically bases the operation of I.C. segmentation on 1) "the identification of the parts of speech and function words" (lexical forms), 2) "marking of the special ties that are signaled by a concordance of forms or by particular intonation contrasts", 3) "identifying the particular arrangement of the Class 1 and Class 2 words that signal the kind of sentence", and 4) "identifying the particular arrangement of the Class 1 words . . . before and after the Class 2 word." In other words, for Fries, I.C. analysis is a segmentation not of a string of morphemes but of a slot and class formula of the sort illustrated above (p.31).

1.3.3 Eugene Nida

"One may also look upon any utterance as consisting of a great series of selections...."

-A Synopsis of English Syntax, p.10

Even though Nida's work appeared before Fries's, it provided a bridge between what we might call the Jespersen-Fries tradition (taxonomic grammar) and the present generative tradition, and, hence, it will be discussed here as transitional between the work of Fries and that of the transformational grammarians to be discussed next.

The major difference between Nida on the one hand and Fries and Jespersen on the other lies in their notions of what form a grammar should take. While Fries and Jespersen were primarily interested in analysis and classification, Nida shifted focus to the way the insights provided by analysis are to be presented. Though he did not develop the notion in great detail, Nida seems to choose a particular format for a grammar because it reflects psychological process by which one produces sentences:

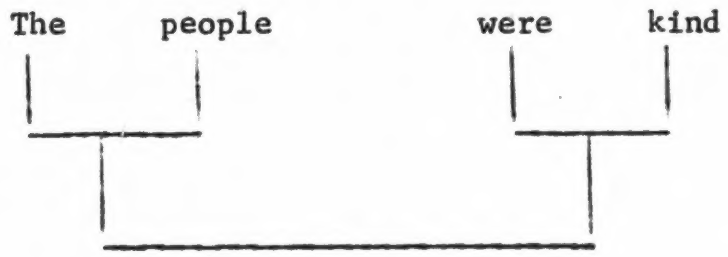
> One may also look upon any utterance as consisting of a great series of selections, if viewed only from the standpoint of the taxeme of selection, and then each section of the outline represents such choices. For example, a person may choose to convey his thought by a major or a minor sentence type. Then in turn he has the choice of an actor-action or a goal-action type. If he chooses the former, he must then choose between transitive, intransitive, or equational types. Having made this choice, he must then choose a subject-head class, and with this come further restrictions as to the possible attributives to the subject expression. Accordingly, as the outline proceeds in depth, the range of choices becomes less. It is this selective basis which is best demonstrated by the outline form. As in the sequence just noted, having chosen a subject expression and determined upon a transitive type

> verb-head expression, the inter-class selection, which implies a selection between function classes of the same coordination value, restricts these verbs to transitive ones, in contrast to intransitive or equational ones, while the intra-class selection, which denotes selections of forms within a class which are selectively determined or conditioned, restricts certain choices of the forms of the verbs so as to be in congruence with subject expressions of person and number. Having listed the verb-head expression, then the various attributives to this must be noted; the first type attributive, which corresponds to what has been in most cases called the indirect object, the second type attributive, which corresponds to the direct object, the alternating attributives, and finally the third attributives, which are generally known as the adverbial attributives. In a similar manner the intransitive and equational clause types are treated, and then the goal-action types, and finally the minor sentence types . . . .

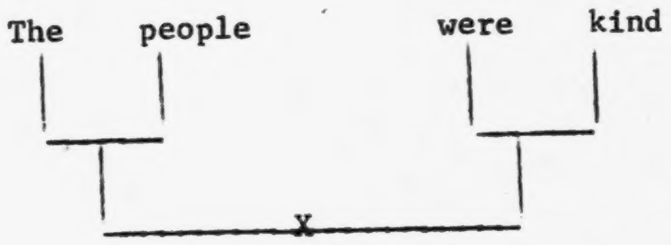
Notice that for Nida this process of increasingly restricted selection begins with a selection of a variety of S (sentence) which is made of a subject and a predicate. Following numerous context sensitive constraints, subject and predicate are selected, as are "attributives" of these functional parts of the sentence. This is a format that has become quite familiar today in the work of Chomsky, Lees, and others.

Nida's outline is a very complex, ordered list of syntactic frames and substitution classes. Each frame specifies five different kinds of relations between constituents; these relations are charted as follows in the later (1960) edition of A Synopsis of English Syntax:

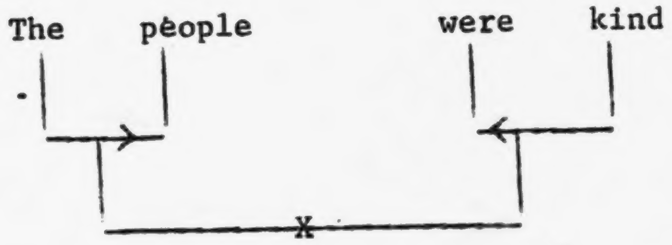
(1) Ranking or layering relationships represented by I.C. diagrams:



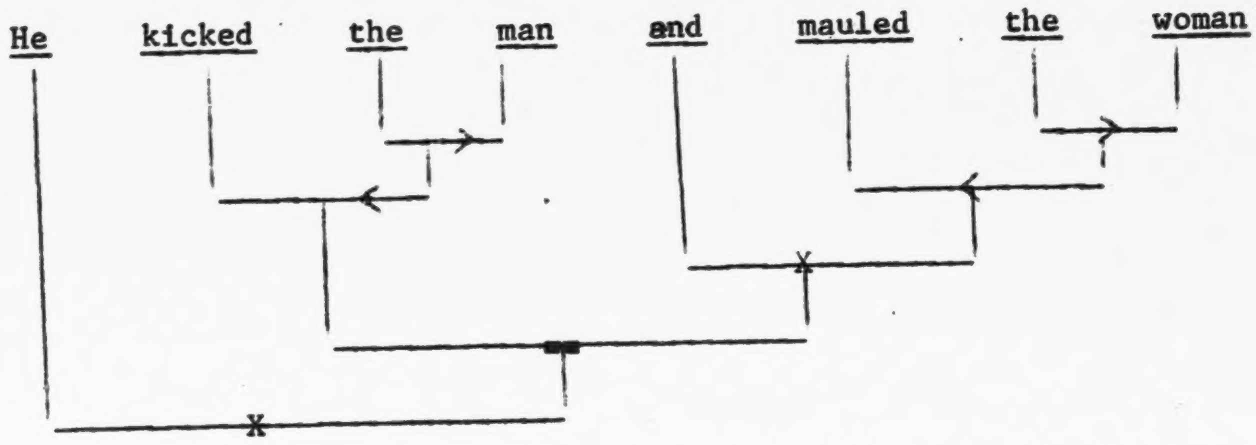
(2) Exocentric relationships, marked by the symbol (x) on the I.C. diagrams:


The
people
were
kind
x

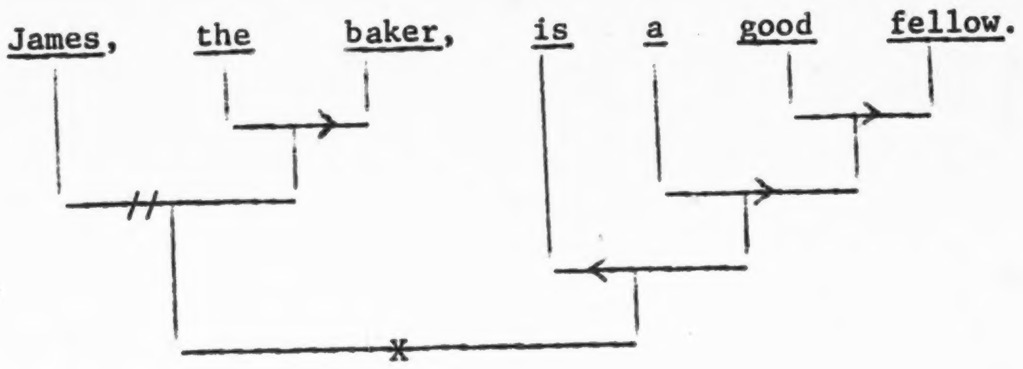
(3) Subordinate endocentric relationships, marked by the symbol (>) or (<) on the I.C. diagrams:


The
people
were
kind
x

(4) Co-ordinate endocentric relationships,[17] marked by the symbol (=) on the I.C. diagrams:


He
kicked
the
man
and
mauled
the
woman
x
=
x

(5) Paratactic relations, marked by the symbol (//) on the I.C. diagrams:


James,
the
baker,
is
a
good
fellow.
//
x

Subject is then defined (p.9) as "the first immediate constituent in the exocentric pattern." Notice that subject is here defined in terms of order (i.e., "the first"), rank (i.e., "the first immediate constituent") and functional relationship (i.e., "in the exocentric pattern"). Subject is clearly a particular grammatical form (aspect A). The grammatical meaning (aspect B) of Subject is only very generally specified in so far as the sentence type selected in the process of producing the sentence is dominated by the labels "Actor-Action" and "Goal-Action", which Nida (p.8) (following Bloomfield) terms "episememes of the tagmemes so designated." Other "episememes" designate substitution sub-classes (aspect C); for example, adjectives are labelled (pp. 43-4) "Qualitative" and "Limiting". What we have designated as aspect D (lexical meaning) is mentioned in the discussion of substitutes, e.g., under the heading personal-definite substitute he, Nida designates (p.47) "Substitution type: Anaphoric and definite identification." However, the relationship of reference to processes like conjoining and pronominalization is not discussed. In describing processes like these, Nida's grammar loses its overall process orientation and becomes exclusively taxonomical. That is, he merely labels the parts of such constructions without specifying selectional constraints on the constituents.

We may surely question whether Nida's grammar reflects the way one produces a sentence; he offers no supporting evidence. And he does not go beyond a surface-level, phrase structure grammar, with its generally recognized limitations (for example, its incomplete description of conjoining). However, this is retrospection, and his over-all goal of writing a grammar which reflects our competence to produce sentences represents a major contribution to English linguistics. A second important contribution is his list of frames and substitution classes for English. They provide a useful check for completeness for anyone who would build on his work.

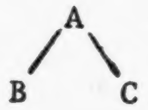
1.3.4 Generative Transformational Grammars

The work of Chomsky, Lees, Fillmore, Klima, and other transformational grammarians has added two profoundly important concepts to the technology of grammatical description; the first of these is the idea that a grammar be generative, that the specification of the analytic reduction of the sentence into its constituents be an explicit, rule-governed process, relying not at all, in principle, on the intuitive analytical powers of the analyst; the second major contribution is use of formal transformation rules which allows strings with endocentric or conjoined interval constructions (with a few exceptions) to be derived from simple strings (basically the simple Subject-Predicate nexus), greatly simplifying the description. We will briefly consider each of these contributions separately and then ask how a generative-transformational grammar treats the four aspects of a grammatical unit which we have posited.

Generative rules can be seen as explicit statements of the sorts of choices Nida describes in the introduction to his A Synopsis of English Syntax. While Nida attempted to present an ordered inventory of the possible constituents of English sentences, Chomsky and others have attempted to describe selection (not production) as a step by step process by which the primitive symbol S (sentence) is expanded into a constituent string. This process in its simplest form is one of substituting an expanded symbol for an unexpanded one:

A ---> B + C

this substitution is also commonly represented in tree diagram or phrase marker:



Note that this sort of substitution is quite different from the substitution in a frame used, for instance, in Fries's work to establish form classes. In a generative grammar substitution constitutes an operational definition of the initial symbol (A). A generative grammar of English is, in theory, an operational definition of the notion "English Sentence". Likewise, Lees (1962, p.8) defines the notion "noun" as follows:

> "Any morpheme (or word, as we wish) which is an expansion of the English grammatical category N in an English grammar, or a transformational replacement of the latter."

A generative grammar of English, then, is not a process for segmenting an English sentence into its parts but an extended, operational definition of "English Sentence" which can be tested according to its completeness and its internal consistency.

The sort of substitution used in a generative grammar is often more complex than the simple expansion,

A ---> B + C

Optional elements can be included in the expansions:

A ---> B + C (D)

Here the parens indicate that D is an optional constituent of the string generated from A. Also, alternative expansion can be indicated by listing alternate constituents vertically between brackets:

$$A \dashrightarrow \left\{\begin{matrix}B\\E\\F\end{matrix}\right\} + C\ (D)$$

Here the initial constituent of the string generated from A is alternatively B or E or F.

Furthermore, special contextual constraints on substitutions can be specified by context sensitive rules:

A ---> B + C/ in the context W...Y

That is, A can be expanded to B + C if the preceding symbol is W and the following symbol is Y.[18] These automatic substitution rules are, in theory, sufficient for generating (i.e., defining operationally) *all the basic sentence types in English.* Each of these sentence types is, in Jespersen's term, a nexus, a Subject-Predicate construction.

However, as transformational grammarians have demonstrated quite convincingly, rules of this sort which expand symbols into branching phrase markers do not describe the conversion of a simple nexus into a new, derived nexus, e.g., a subordinate clause, a compound sentence, etc. Here we hark back to Jespersen's observation of the agnate relation of

(a) the barking dogs

(b) The dogs bark.

Transformation rules are rules which among other things derive the endocentric construction (a) from the underlying exocentric nexus (b), hence greatly simplifying the description and accounting formally for the intuition that (a) and (b) are related in some deep way.

Lees and others have shown that it is possible to describe nearly all so-called modifier-head junctions or endocentric relations via transformation rules. As Lees demonstrates (1960a), all the subordinate and coordinate constituents of the noun phrase (except the determiner system, which is probably not subordinate, and some adjectives like major, few, poor, etc., can be derived from underlying sentences.

From this distinction between simple (or kernel) constructions and transformationally derived constructions has come the very important contrast between surface and deep sentence structure.[19] Any grammar which does not explicitly (i.e., formally) describe the relation of a noun phrase like (a) to an underlying sentence (b) is clearly incomplete. Part of the grammatical meaning of (a) is our recognition of the action-actor relationship of barking to dogs. Merely to label barking as a modifier of dogs manifested by the present participle of an intransitive verb is to describe only the surface structure of the construction.

Questions about the formal relation of deep to surface structure, particularly the question of the sort of information a grammar should provide as a basis for both lexical and phonological interpretation, has led Chomsky recently into very profound revisions of generative transformation grammar. In _Aspects of the Theory of Syntax_ (p.64) Chomsky asks how information of the sort provided by traditional grammar "can be formally presented in a structural description and how such structural descriptions can be generated by a system of explicit rules." He lists three sorts of information that traditional grammars provided:

(1) _Lexical form classes_; e.g., in the sentence, "Sincerity may frighten the boy", "_frighten the boy_ is a verb phrase (VP) consisting of the verb (V) _frighten_ and the noun phrase (NP) _the boy_; _sincerity_ is also a NP; the NP _the boy_ consists of the determiner (Det) _the_, followed by a noun (N); the NP _sincerity_ consists of just an N; _the_ is, furthermore, an article (Art); _may_ is a verbal auxiliary (Aux) and, furthermore, a modal (M)." This is the sort of information provided by Fries in _The Structure of English_.

(2) _Grammatical forms (or relations)_. The NP _sincerity_ functions as the Subject of the sentence, whereas the VP _frighten the boy_ functions as the Predicate of this sentence; the NP _the boy_ functions as the Object of the VP, and the V _frighten_ as its Main Verb; the grammatical relation Subject-Verb holds of the pair (_sincerity_, _frighten_), and the grammatical relation Verb-Object holds of the pair (_frighten_, _the boy_).

(3) _Lexical subclasses_. The N _boy_ is a Count Noun (as distinct from the Mass Noun _butter_ and the Abstract Noun _sincerity_) and a

Common Noun (as distinct from the Proper Noun John and the Pronoun it); it is, furthermore, an Animate Noun (as distinct from book) and a Human Noun (as distinct from bee); frighten is a Transitive Verb (as distinct from occur); and one that does not freely permit Object deletion (as distinct from read, eat); it takes Progressive Aspect freely (as distinct from know, own); it allows Abstract Subjects (as distinct from eat, admire) and Human Objects (as distinct from read, wear).

Chomsky states (p.64) that the information in (1), (2), and (3) above is "substantially correct and is essential to any account of how the language is used or acquired." As his discussion of the formulation and generation of this information is so close to the problem raised in this chapter (i.e., how the four aspects of grammatical units are treated in a grammar), we shall examine his discussion in detail. We shall assume a fairly close correspondence between (1), (2), and (3) and aspects (C), (A), and (D) correspondingly, as our labels on Chomsky's examples suggest.

The information in (1) above can be presented by a labeled bracketing or tree diagram (Phrase-marker) and generated by the simple sorts of rules described earlier. What we have called aspect C is easily accounted for in a generative-transformational grammar.

The sorts of information in (2) above (our aspect A), Chomsky argues (p.69), "are already represented in the Phrase-marker, and no new rewriting rules are required to introduce them." He defines the notions subject, predicate, direct object and main-verb as follows:

Subject - of: [NP, S]

Predicate - of: [VP, S]

Direct-Object-of: [NP, VP]

Main - Verb - of: [V, VP]

That is, the subject of S is the NP dominated by S in the Phrase-marker, the object of S is the NP dominated by VP in the Phrase-marker, etc.[20] Chomsky believes that by thinking of these relations (subject, object, etc.) as categories we make the mistake of trying to label them in surface structure. For example, in the sentence (Chomsky, 1965, p.70):

John was persuaded by Bill to leave.

"John is simultaneously object-of persuade (to leave) and subject-of leave." Chomsky goes on to distinguish the familiar "logical" (i.e., deep) and grammatical (i.e., surface) subjects of sentences.

Chomsky limits the ascribing of function notions to what he calls the major categories (i.e., a lexical form class that dominates a string in a Phrase-marker). These major categories are probably the categories noun, verb, adjective, and adverb.[21] It is not clear how or whether Chomsky would define such traditional grammatical relations as indirect object, object complement, and, especially the various functions (instrumental, accompaniment, location, time, etc.) that are all, according to his formulation, [Adv, S].

It is when he comes to describing and generating the information in (3) above that Chomsky proposes a radical change in the transformational model, a change which forces reconsideration of his accounting for lexical form classes and grammatical forms as discussed above.

In order for the grammar not to generate sentences like,

*The boy may frighten sincerity.

there must be sub-classification of nouns. (Because this subclassi- is based entirely on the reference of the class of lexical items--features like human, abstract, common, proper, etc.--it is clearly aspect D in our system.) Chomsky proposes an elaboration of the traditional technique of substitution in frames to establish these subclasses purely syntactically. Sub-classes of verbs are to be defined in terms of co-occurrence possibilities with subject and object.[22] This requires that subjects and objects carry information like: ± Common, ± Count, ± Animate, ± Human, ± Abstract, etc. The complex symbols are not, however, representable in tree diagrams with nodes labeled by symbols for lexical classes.[23]

At this point, symbols in a generative grammar carry through the generative process three kinds of information, e.g.,

Object /N/ -Count, + Common, + Human

represents boy in the sentence

Sincerity may frighten the boy.

Clearly, something like our aspects, A, C, and D are represented here,[24] and if the suggestions of Fillmore are accepted, B will be added.

Fillmore (1967) argues that each NP in the grammar is dominated by K (kasus) and that these cases show the "deepest" deep structure. Thus, in the following sentences there is an underlying system of grammatical relationships deeper than subject and object:

The door will open.

The janitor will open the door.

The janitor will open the door with the key.

This key will open the door.

The door will open with this key.

The door will be opened with this key.

The door will be opened by the janitor.

Underlying these sentences are the categories agentive (*the janitor*), instrumental (*this key*), and ergative (*the door*), plus a verb class that can be symbolized [+ergative, ± agentive, ± instrumental]. In Fillmore's system, unlike Chomsky's, it is not subject and object which select the verb (these are surface structure relations) but the *cases* of the major categories (except the verb). We shall have frequent cause to mention Fillmore's work below. Suffice it to say here that if we add Fillmore's *case* to Chomsky's complex grammatical unit, we get all our four aspects:

	Grammar	Lexicon
Form	e.g., Object	e.g., Noun
Meaning	e.g., Ergative	e.g., -Count, +Common, +Animate, +Human

Chapter II

A GENERATIVE TAGMEMIC GRAMMAR

2.1 Longacre's Generative Model of Tagmemics

In the introduction to his recent book, Grammar Discovery Procedures, Robert Longacre sketches a generative formulation of a tagmemic grammar, attempting to make explicit the previously implicit generative power of tagmemic grammars.[1] This model will be examined in some detail as the basis on which our description of the English subject will be built. First the model as Longacre introduced it will be described; then some problems in Longacre's formulation will be discussed; finally, at the end of this chapter, a modification of Longacre's model will be presented.

Longacre's generative process assumes a completed taxonomic (phrase structure) grammar in which (1) all string-constituent constructions are given, arranged according to levels (e.g. sentence level, clause level, phrase level, etc.), and (2) both slots and fillers of the constituents are given (e.g. Subject: phrase type 1, phrase type 2, clause type 1, clause type 2, etc.). It is not clear from Longacre's brief formulation how this inventory of formulae is to be presented in the grammar, but there is some suggestion (1964b, p.14-15) that they can be ordered into matrices. This base grammar might be of the form:

Clause type 1 = +A: s/t +B:m/n ± C:s/t/u
Clause type 2 = +A: s/t/u +E: m/o ±C:s/t/u
Etc.

These formulae are to be read, "Clause type 1: obligatory slot A filled either by syntagmeme s or t, plus obligatory slot B filled either by syntagmeme m or n, plus optional slot C filled either by syntagmeme s, t, or u." Then formulae for phrase syntagmemes s,t,m,n,o,u can be given. Finally word level formulae will be listed.

Given an inventory of constructions of this type--though a great deal more complex, Longacre's generative process consists of three kinds of operations on these formulae: a reading operation (R), a permutation (or re-ordering) operation (P), and an exponence operation (E). In the reading operation, strings are formed including just the simple slot designations (the symbols preceding the colons); alternate readings are possible if the formula includes optional constituents. Thus for Clause type 1 above, there are two possible readings: A B and A B C.[2] In the permutation operation, a particular reading (e.g. A B C) is permuted in all possible ways. Suppose, for instance, that A B C can be permuted in three ways, including identity permutation: A B C (identity), C A B, and A C B. Next, in the exponence operation, there are two stages: first, the symbols in each permuted reading are replaced by labels for particular exponents or fillers, e.g. A B C is rewritten s m t; then, the labels for the fillers are replaced by formulae for filler syntagmemes, e.g. s is replaced by the formula +G + H ± L (perhaps a noun phrase construction). Then these operations are carried out on this phrase level formula, and so on to word level, until only particular morphemes or labels for morpheme classes remain.

Let us now attempt to formalize these rules further, using the conventional arrow to mean "replace string to left of arrow by string to right of arrow," i.e. $A \rightarrow +B+C$. As an illustration, we shall use the Trique intransitive clause which Longacre describes in Grammar Discovery Procedures: $+P_i + S \pm L \pm T$ (P_i= intransitive predicate, S = subject, L = locational, T = temporal). The first rule is a set of alternative readings of the formula, and it might be written:

$$(1) \quad +P_i + S \pm L \pm T \rightarrow \begin{bmatrix} PS \\ PSL \\ PST \\ PSLT \end{bmatrix}$$

This is a reading rule which generates all the possible readings of the Trique clause formula; taking only one of the four readings, we can carry out permutations according to two rules which Longacre states (1964a, p.27) as, A) "One symbol at a time may be moved to the left of P" and B) "Contiguous LT may vary to TL." These rules may be represented formally as:

$$(2) \quad PSLT \rightarrow \begin{bmatrix} PSLT \\ SPLT \\ LPST \\ TPSL \\ PSTL \\ SPTL \end{bmatrix}$$

In a complete grammar, permutations for each reading would be given, or else a general statement of rules like A and B above. These would both be context sensitive rules, perhaps of the sort, (2a) $\ldots x \rightarrow \#x\ldots / \ \#P\ldots x\ldots$ (or $\#\_P\ldots$)

That is, in the Trique intransitive clause, any non-initial symbol may be permuted to initial position if and only if the context is P initial. Choosing a particular permutation, we then have the exponence rule,

(3) $\mathrm{PSLT} \rightarrow \mathrm{Ph}_{1_i}\,\mathrm{Ph}_{11}\,\mathrm{Ph}_{41_1}\,\mathrm{Ph}_{31}$

Actually, as Longacre tells us, P has six exponents, S has eleven, L has six, and T has seven. Here a particular exponent (or filler) is given for each of the symbols. Presumably alternant fillers for each tagmeme are stored in the base inventory of the grammar, or else alternants could be listed in brackets:

(3a) $$\mathrm{PSLT} \rightarrow \begin{bmatrix} \mathrm{Ph}_{1_i} \\ \mathrm{Ph}_{2_i} \\ \mathrm{Ph}_{3\cdot 3} \\ \mathrm{Ph}_{4} \\ \mathrm{Ph}_{5} \\ \mathrm{Ph}_{6_i} \end{bmatrix} \begin{bmatrix} \mathrm{Ph}_{11} \\ \ldots \\ \ldots \end{bmatrix} \quad \text{Etc.}$$

Longacre does not list all of the exponents for each tagmeme, nor does he discuss the possibility of collocational constraints in the selection of exponents, which would necessitate context sensitive rules and possibly the ordering of exponent selection. He does state that L and T may have as exponents subordinate clauses, and that S has an exponent which is a relative sentence; both examples of cross-level embedding.

In the exponence rule (3) above, Ph_{1_i} = active qualifier main verb phrase of intransitive class, Ph_{11}= qualifier noun phrase, Ph_{41_1} = relator axis phrase of locational class, and Ph_{31} = ordinal temporal phrase. Hence, all clause level tagmemes (P,S,L, and T) are neatly filled by phrase level syntagmemes.

We now substitute phrase level syntagmeme formulae for the labels in (3), referring to the base inventory for these formulae (1964a,p.29):

(4) $Ph_{1_1} Ph_{11} Ph_{41_1}$ Ph_{31} → ($\pm$adv$\pm$Aux +Mn $\pm$Md $\pm$r) ($\pm$i $\pm$q +h $\pm$A $\pm$d) (+R +Ax)

(+U +Or)

If we carried out exponence operations for all of the permutations of all of the readings of the clause syntagmeme $+P_1$ +S $\pm$L $\pm$T, we could derive a very large number of exponential combinations, "somewhere around 200,000" according to Longacre's calculations.

Now that the operations have been carried out on the initial reading of the clause level syntagmeme, we repeat the reading, permutation, and exponence operations at phrase level, i.e. on the formulae in (4), as follows:

(5) ($\pm$adv $\pm$Aux +Mn $\pm$Md $\pm$r) → adv Aux Mn

(6) ($\pm$i $\pm$q +H $\pm$A $\pm$d) → H A d

(7) (+R +Ax) → R Ax

(8) (+U +Or) → U Or

Rules (5) -- (8) represent particular readings of the formulae for phrase level syntagmemes in (4). As there are only identity permutations for Trique phrase level syntagmemes, we will not list them here but rather proceed directly to the first exponence operation on these readings.

(9) adv Aux Mn → a^5ʔ ('already') W_5aux W_5

(10) H A d → Noun Adjective dą^{3}h ('a certain')

(11) R Ax → ri^{3}ki^3 ('under') Ph_{11}

(12) U Or → gwi^3 ('day') Ph_{61}

We note here that actual morphemes have entered the derivation. These are presumably lexical markers of particular Trique phrase level syntagmemes, not random choices from morpheme classes, though Longacre does not so state. Other exponents of phrase level tagmemes are labels for word level syntagmemes (e.g. W_5 in rule 9) or labels for word classes (e.g. Noun in rule 10) or embedded phrase level syntagmemes (e.g. Ph_{11} in rule 11). We now proceed to the second exponence rule, substituting formulae for the labels of syntagmemes:

(13) W_5aux → (+asp +mood +vs')

(14) W_5 → (+asp +mood +vs')

(15) Ph_{11} → (±i ±q +H ±A ±d)

(16) Ph_{61} → (±m +sc ±int)

We now cycle back to readings of the formulae of (13) - (16):

(17) (+asp +mood +vs'$_{aux}$) → asp mood vs'$_{aux}$

(18) (+asp +mood +vs') → asp mood vs'

(19) (±i ±q +H ±A ±d) → H d

(20) (±m +sc ±int) → sc int

As there are also no permutations possible here, we will give exponents for the readings (17) - (20); according to Longacre, all symbols to the right of arrows in these readings (17-20) are labels of morpheme classes except those in (19). The exponence rule for (19) is:

(21) H d → Noun dą^{3}h

We may now substitute the results of rules (5) - (21) into the string (4) to get the terminal grammatical string. (We will leave in parentheses the terminal representations of the tagmemes in the initial clause level syntagmeme +P$_i$ +S ±L ±T.)

Terminal string = (a^{5}ʔ asp-mood-vs'$_{aux}$asp-mood-vs') (Noun Adjective dą^{3}h)
(ri^{3}ki^3 Noun dą^{3}h) (gwi^3 sc int)

Then, according to Longacre, lexical substitution from a cross reference dictionary yields the clause,

(a^{5}ʔgą^{3}ʔą^{34}h ga^{5}wi^{5}ʔ) (žu^{3}we^3 ga^{5}ci^5 dą^{3}h) (ri^{3}ki^3 čų^3 dą^{3}h) (gwi^{3}go^{4}gą^{5}ʔą^{3}h)

'(Already went to die)(that white dog) (under that box) (the twenty-fourth day)'.

At this point the lexical string must be given proper phonological structure from a "semiautonomous" cross reference phonology in order to generate the terminal representation of a Trique clause. Longacre does not describe the structure of either the cross reference lexicon or the cross reference phonology.

Aside from some problems concerning the general structure of the model, several particular questions arise here. We have noted some of them in passing. How is context sensitivity described? In rules (13) and (14), for instance, there must be some concord between aspect-mood in W$_5$aux and in W$_5$. Longacre states (p.30) only, "Indications of aspect and mood will be supplied according to particular allomorphs of aspectual and modal morphemes as given in the dictionary." If there is some agreement between subject and predicate, how is this to be described? Also, rule (15) seems a needless repetition, as the information has already been given in rule (4). Other rules might also be generalized; for example, the permutation rule might be stated in a general way rather than by a list as in rule (2). Furthermore, Longacre seems to have overlooked the major problem, discussed above in relation to transformational grammars, of the sub-classification of lexical classes. It seems doubtful, for instance, that in rules (10) and (21) any Noun (or Adjective) may be chosen.

These and other particular problems arise when we study Longacre's fragment of Trique grammar. Some of them demand a knowledge of Trique to be resolved and hence must remain unresolved here. Others question the adequacy of the model as it has been briefly presented. We can better examine these latter questions by attempting to apply Longacre's model to English. Is tagmemics inadequate as a modern, generative grammar, as some have claimed[3], or can Longacre's attempt to formalize the implicit generative capacity of tagmemics provide the framework for a generative tagmemic grammar of English?[4]

2.2 Some Problems in Longacre's Model

We will try to show that the major problems in Longacre's model are 1) that he does not provide enough information about tagmemes to allow a complete description of certain features of English, and 2) that he fails to distinguish surface structure and deep structure and hence obscures many important grammatical insights. However, we will also attempt to demonstrate that his model can be adapted so as to overcome these problems. Both of these problems can be revealed if we attempt to describe clause level conjoining in English. The discussion will also serve to demonstrate the relevance of all the four aspects of the tagmeme described earlier in Chapter I section 1.1

In Syntactic Structures (p.36) Noam Chomsky introduces the description of conjoining as follows:

> ...the possibility of conjunction offers one of the best criteria for the initial determination of phrase structure. We can simplify the description of conjunction if we try to set up constituents in such a way that the following rule will hold:
>
> If S_1 and S_2 are grammatical sentences, where S_1 differs from S_2 only in that X appears in S_1 where Y appears in S_2 ... and X and Y are constituents of the same type in S_1 and S_2, respectively, then S_3 is a sentence, where S_3 is the result of replacing X by X + and +Y in S_1....

Chomsky goes on to argue, in substance, that a phrase structure grammar (as he defines it) has no way of incorporating such a rule--that is, a rule which describes a binary transform. And because this rule leads to considerable simplification in writing a grammar, "it provides," he writes (p.38), "one of the best criteria for determining how to set up consitiuents."

We will argue here that Chomsky was right in giving great importance to the description of conjoining, that it does indeed put a number of important constraints on the way we describe the constituents of a grammatical string, and that, therefore, the specification of tagmemic conjoining rules will be both interesting in allowing us to compare and evaluate grammatical theories and necessary in establishing criteria which a grammatical description must meet.

Why is the description of conjoining of such importance in writing a grammar? There are two main reasons:

1) Conjoining, like embedding, is theoretically an open-ended (recursive) operation. Although there are undoubtedly psychological and rhetorical constraints on all recursive rules, there seems no reason why, given a particular string like:

John, Bill, and Frank are here.

we cannot always add one more item to the list:

John, Bill, Frank, and Larry are here.

Furthermore, this recursiveness is not the same operation as recursive embedding; that is, coordination is different from subordination. Conjoining is an operation repeating constituents of a construction (syntagmeme) at a particular level (e.g. sentence, clause, phrase, word), while subjoining, on the other hand, is an operation not of repeating but of complementing a particular constituent in a construction. Conjoining adds to the number of constituents at a level, subjoining does not. It is important not to confuse these two kinds of recursiveness if we wish a grammar to reveal not just the form but also the relationship of constituents in a string. Conjoining, then, requires that a grammar be able to describe what I will call linear recursiveness.

2) The second reason why conjoining is an important constraint on the form of a grammar is that it is an especially context sensitive operation. As Chomsky observed, conjoined constituents must be "constituents of the same type." The difficulty lies in saying what "of the same type" means and in building into the grammar a means of marking such constituents. Clearly, the lexical form of the conjoined elements does not determine whether or

not they may be conjoined. That is, class labels like Noun Phrase, Prepositional Phrase, Adverb, etc. do not provide sufficient information about the constituents of a string to allow us to write conjoining rules. Conjoined constituents can have different lexical forms:

John stepped into the water carefully and without a word.

And constituents of the same lexical form cannot always be conjoined:

*John danced with Mary and with a limp.

Furthermore, if we define "of the same type" as "dominated by the same node in a phrase structure tree," we cannot easily explain such conjoinings as

I floated and he swam across the lake.[5]

Also, conjoining or coordination is a label for at least three different linearly recursive operations: conjoining proper, disjoining, and alternating. That is, conjoining with and is different from conjoining with but and or, as the following three sentences demonstrate:

John danced with Mary and with Sue, too.

John danced with Mary or with Sue (*too).

*John danced with Mary but with Sue.

Conjoining clearly requires more information about constituents of a construction than can be derived from a tree diagram. Conjoinable constituents are not necessarily those of the same lexical form or those dominated by the same node in a phrase structure tree.

Can Longacre's rules describe linear recursiveness, particularly conjoining? In his reading rules Longacre includes the possibility of superscripts on tagmeme symbols; these superscripts were left out of our presentation of his model above, both because it made the description of his model simpler and because we wished to give special attention to them here. In his original presentation, two tagmemes in the Trique intransitive clause syntagmeme bore superscripts, indicating that these tagmemes could be optionally repeated in a reading. Thus in rule (1) there are actually nine rather than four possible readings:

$$(1)\ +P_i\ +S\ \pm L^2\ \pm T^2 \qquad \begin{bmatrix} \mathrm{PS} \\ \mathrm{PSL} \\ \mathrm{PST} \\ \mathrm{PSLT} \\ \mathrm{PSLL} \\ \mathrm{PSTT} \\ \mathrm{PSLLT} \\ \mathrm{PSLTT} \\ \mathrm{PSLLTT} \end{bmatrix}$$

This possibility of optionally repeating a tagmeme is motivated by the need for linear recursiveness. In fact, repeating a tagmeme in a reading rule is equivalent to conjoining, for if a tagmeme is actually repeated, then, at least in English, a conjunction marker (and, but, or, and others, including special intonation) is obligatory. For example, if any of the tagmemes in the following syntagmeme are repeated, a conjunction marker is necessary:

John fishes with a fly rod.

$+S\ +Pr_i \pm Inst.$

If $+S^2$ is read,we get

John and Bill fish with a fly rod

If $+Pr_i^{\ 2}$ is read, we get

John fishes and hunts (?) with a fly rod.

If $+Ist^2$is read, we get

John fishes with a fly rod and with a casting rod.

In none of these instances can we delete the conjunction marker if we repeat a tagmeme. In those cases where a tagmeme seems to be repeated we would argue that there are actually two different tagmemes present. For example, in the following sentence the Location tagmeme appears to be repeated:

I live at 2165 Newport in Ann Arbor.

Here the proper analysis probably requires two different location tagmemes, perhaps $\mathrm{Loc}_{\mathrm{area}}$ and $\mathrm{Loc}_{\mathrm{point}}$. The evidence is that the two different location tagmemes can be permuted independently here:

In Ann Arbor I live at 2165 Newport.

(though not *At 2165 Newport I live in Ann Arbor)

In the case of repeated or conjoined tagmemes, the two must be permuted as a unit:

John bathes in the morning and at night.

In the morning and at night John bathes.

*At night John bathes in the morning.

Though Longacre's superscripts can describe some instances of conjoining, there are other instances which they cannot at present describe, for sequences of tagmemes as well as single tagmemes can be conjoined:

John runs slowly and walks quickly.

I float and John swims across the lake.

John walks quickly in the morning and slowly in the afternoon.

Because conjoining or repeating involves both single tagmemes and sequences of tagmemes, superscripts on single tagmemes are clearly inadequate.

Hence the first change in Longacres model is proposed, motivated by the need for linear recursiveness. Following the reading operation and the permutation operation, an optional conjoining operation (K) is introduced. It appears that in English this operation is somewhat different at different levels of the grammar. For example, conjoining with but (more correctly, disjoining) operates under different constraints at the clause and phrase levels. At clause level, the operation seems to require at least two differences between the conjoined constituents, one of them often (but not always) a negative.[6] For example, a grammar should generate only the first two of the following three sentences:

John bathes in the morning but not at night.

John bathes in the morning but showers at night.

*John bathes in the morning but at night.

This same constraint does not seem to apply to phrase level conjoining, at least not in the same way. There are not two overt structural differences between the conjoined constituents in the following phrases:

A small but tough boy.

A simple but haunting melody.

We do feel a semantic contrast between the conjoined elements in these two phrases. Smallness seems to imply some negation of toughness, and this is just the point emphasized by but. The fact that the negative sense is in some cases marked and in some cases unmarked means that a grammar will have to carry some semantic information into grammatical description. That is, the description of conjoining with but will indicate the need for semantic contrast between items selected to interpret lexically the grammatical representation of the syntagmeme. Some degree of lexical meaning (Aspect D of our proposed four part tagmeme) is relevant for grammatical description.

If conjoining is different at different levels, it may well serve as a diagnostic to determine levels in a particular language. As we shall demonstrate below, the possibility of conjoining constituents serves in many ways as a major criterion for, in Chomsky's words, "determining how to set up constituents."

It seems clear that the conjoining rules should follow permutation rules and precede exponence rules. As we saw above, conjoined tagmemes are permuted (if they can be permuted)as a unit. Rather than build this requirement into all permutation rules, it seems simpler to have tagmemes permuted and then conjoined. Conjoining must precede exponence because of several constraints that conjoining makes on the selection of fillers for tagmemes, including the requirement for semantic contrast in conjoining with but which was discussed above.

Considering only clause level conjoining and only conjoining, with and (in order to simplify the discussion for the moment), we might set up a rule that any tagmeme or sequence of tagmemes in a permuted reading of a syntagmeme may optionally be repeated. This is not, however, a sufficiently precise rule to include all of the constraints on conjoining even with and. Longacre's reading rules provide information only about the grammatical relation or form of tagmemes (e.g. subject, predicate, object, etc.). Two tagmemes may have the same grammatical form and still not be conjoinable, as in the following examples:

John hit Bill with a stone.

A stone hit Bill.

*John and a stone hit Bill.

Here both John and a stone are subjects of the verb hit, but they are not conjoinable. More than just the relational symbol subject (S) must be given to prevent the conjoining of subjects like John and the stone, which is another way of stating what we said earlier; that is, that conjoining requires more information about constituents than can be derived from a labeled tree diagram.

Hence, there is a deficiency in the sorts of units Longacre uses as constituents of his string formulae. As we saw in Chapter I, there are four kinds of information about units that are relevant to grammatical description: grammatical form and meaning, and lexical form and meaning. Longacre's formulae initially provide only specification of grammatical form, e.g. subject, predicate, object, etc. Later, in the first exponence rule, Longacre's model provides specification of lexical form, e.g. Noun Phrase$_n$, Adjective Phrase$_n$, etc. Throughout his grammar only two aspects of the tagmeme are operative. These two aspects of the tagmeme (grammatical slot and lexical filler class) mark only the surface structure of syntagmemes, omitting features of deep structure that are necessary to describe grammatical phenomena of which conjoining is only one instance.

The difference between deep and surface structure has often been obscured in tagmemic grammars, though the distinction is implicit throughout Pike's work. Pike frequently has designated a tagmeme in terms of both grammatical form and grammtical meaning with such notations as "subject-as-actor", "subject-as-goal", etc. It is important to realize the relevance of this traditional designation of the grammatical unit. For instance, in the following sentences grammatical meaning is invariant while grammatical form changes:

Subject as actor: John feeds the cat in the morning.
Subject as goal: The cat is fed by John in the morning.
Subject as time: In the morning is when John feeds the cat.
Subject as action: Feeding is done in the morning.
(or, more awkwardly: The feeding of the cat is done by John in the morning.)

These sentences illustrate a surface reordering of underlying categories of grammatical meaning, i.e. actor action goal time. The simple designation of the constituents of these sentences as grammatical function (here, grammatical form) plus filler class (here, lexical form) obscure this invariance between them. That is, Longacre's notation (e.g. Subject: Noun Phrase) identifies only the surface structure of the sentences.

Furthermore, labels for surface relations (grammatical form) and labels for underlying categories (grammatical meaning) have often been mixed or represented as a single system in tagmemic grammars. For instance, Longacre gives the formula for the Trique intransitive clause syntagmeme as +S(ubject) + P(redicate)$_i$ ± L(ocational) ± T(emporal). There are two basically different kinds of designations in this formula: S and P represent grammatical relations (grammatical forms) while L and T represent semantic categories (grammatical meanings). As we have seen, the temporal element may appear in the surface structure as subject:

In the morning is when John feeds the cat.

The locational element may also appear as subject:

Over the fence is an automatic double.

Locational and temporal may also appear in the predicate:

John boxes flowers. (John puts flowers in boxes.)

John winters in Florida. (John lives in Florida during the winter.)

It seems clear that there are two different kinds of units in Longacre's syntagmeme formulae. These correspond to aspects A and B of our unit described in Chapter I. Aspect A includes relations like subject, predicate, object, complement, modifier, head, etc. Aspect B includes semantic categories (somewhat similar to cases) like actor, goal, instrument, time, location, manner, and many more. As noted earlier, Jespersen and Fries make a very similar distinction between grammatical subject and notional (elsewhere logical) subject in, for example, the sentence,

He happened to fall.

where "he" is the grammatical subject of "happened to fall" but also the notional subject of "to fall" and "he-to-fall" is the notional subject of "happened". In our terms, grammatical relations represent constituents of the surface structure, and notional relations represent constituents of the deep or underlying structure.

Returning to the problem of conjoining, if we specify that tagmemes must be equivalent in both grammatical form and grammatical meaning, such strings as the following will not be generated:

*John and the stone hit Bill.

Here "John" would be marked by the complex symbol $\text{Subject}_{\text{actor}}$ and "the stone" by $\text{Subject}_{\text{instrument}}$. Our conjoining rule would now contain the added stipulation: tagmemes are K-equivalent (conjoinable) if and only if their grammatical function and their grammatical meaning are the same. It will be demonstrated later how the information for such context sensitive rules can be built into Longacre's reading rules.

However, still further information about tagmemes is necessary to describe K-equivalence. It is clear that the rules should not generate a string like,

*I and I are going.

Likewise, it seems clear that the subject of the following sentence refers to two different people:

John and John are going.

This information is important in a grammar in the same way that it is grammatically relevant to identify the single referent of the subject and the complement in the sentence

That man is my brother.

The pronoun substitute here is "he", not "they." This information about reference is provided by aspect D (lexical meaning) of the grammatical unit. The relevance of aspect D in writing pronominalization rules has already been pointed out. For example, if the two instances of "John" in the following sentence have the same reference (lexical meaning), then reflexive pronominalization is obligatory:

John hit himself in the foot. John hit John in the foot.

Hence, it seems necessary that information about lexical meaning be included in our notation, and that the definition of K-equivalence requires this information. The rule can now be rewritten:

> Any tagmeme or sequence of tagmemes in a permuted reading of a syntagmeme may optionally be repeated if and only if their grammatical form and grammatical meaning are the same and their lexical meanings are different.

In Summary, considering only clause level conjoining with and, we have seen the necessity for three basic changes in Longacre's model:

1) At least one other kind of rule, a conjoining rule, must be added to Longacre's model.

2) Tagmemes must distinguish grammatical form from grammatical meaning and carry information about both.

3) The grammar must provide information about lexical meaning, at least to the extent of indicating whether tagmemes in a syntagmeme have "some" or "different" reference.

In the following section we shall describe how some of these conditions can be met in a generative tagmemic grammar.

2.3 Deep and Surface Structure in Tagmemic Reading Rules

In describing a new format for tagmemic reading rules, motivated, as we have seen, by an established need for greater context sensitivity and a clear separation between surface forms and underlying semantic categories, we shall discuss first only the rather gross generative capacity of the model. There seem to be two useful sorts of generative grammars, those that present directions for re-write rules and those which present these rules themselves in detail. Of course, no grammar of the latter sort has yet been written: we have to be content with rather indelicate fragments of grammars. A complete tagmemic grammar will have to provide all possible exponence rules of all possible permutations of all possible readings of English constructions--a very long grammar even if we begin with clause level formulae.[7] In the pages and chapters which follow, we will try to become progressively more precise in the formulation of the grammar, quite aware, however, that we will not reach that level of delicacy where exceptions may no longer be found.

Let us begin with a very simple formula for an English intransitive clause, as it might be presented in Longacre's generative model: +S(ubject) +P(redicate) ±L(ocation) ±T(ime). We have already seen that a formula of this sort mixes symbols for two different aspects of tagmemes, grammatical forms (surface relations) and categories of grammatical meaning (deep semantic relations). Thus we must split the formula above into these two aspects:

A = + S + P ____ ____

B = ____ ____ Loc. Time

To supply labels for the other two categories in B is not too difficult. If we let the formula be illustrated by the sentence,

John runs in the park in the morning

then we can label these categories of B Agent (which seems a more general term than actor) and Motion, giving us the underlying string of grammatical categories,

Agent Motion Location Time

We shall delay until later the justification of these categories; each one must have grammatical relevance.

To identify the two missing grammatical relations in A above is much more difficult. What is the grammatical function of "in the park" and "in the morning"? They are optional, non-nuclear constituents of the syntagmeme, frequently called sentence adverbials.

Interpreted as a lexical class, adverbials have defied definition. The class of words which we might label "manner adverbials" is clearly equivalent to a sub-set of the class "adjectives." Lees (1962, pp. 16-17) and others have pointed out the relation of manner adverbials with "modifiers" in nominalized verb phrases:

John drives safely	John's safe driving
John explains it clearly	John's clear explanation of it
John is awfully	John's awful silence

Here aspects A, B, and C of the constituents must be distinguished. We can distinguish three lexical classes (C): Nouns (John, it), Verbs (drive, explain, be), and a third class with no traditional name;[8] for the present label them Adjectives (safe, clear, silent, awful). Let us say that these Adjectives manifest the semantic category (B) "manner" in the first two examples, "degree" in the third. Their grammatical form is different depending upon the surface grammatical relationship. Notice, however, that the change of grammatical form in each case is accompanied by a shift of level. A clause level adjunct ("safely" in "John drives safely") becomes a phrase level subjunct ("safe" in "John's safe driving"). The terms "adjunct" and "subjunct" are deliberately contrasted here. "Safe" modifies "driving" in a headed construction (Jespersen's junction), while "safely" cannot be said to modify a single term but rather the whole nucleus of the syntagmeme (John-drives). Let us therefore call the grammatical relation of these so-called "sentence adverbials" adjunct.

In English clauses there seem to be a number of rather different adjuncts. For instance, some can be permuted to sentence initial position, others cannot:

In the morning John runs in the park.

*In the park John runs in the morning.

These differences in surface form seem clearly differentiated by aspect B of the tagmeme, however. Here, for instance, we can say that Adjunct (Time) can be permuted to sentence initial position, while Adjunct (location) can not. It seems sufficient to mark the grammatical relation (aspect A) of these tagmemes with a single symbol, therefore: Ad 'adjunct'.

We can now distinguish the two kinds of structure represented by the undifferentiated formula $+S +P_i \pm L \pm T$:

$A = +S +P \pm Ad \pm Ad$

$B = Ag\ Mo\ Loc\ Temp$

A Reading rule now becomes a reading of the surface form (A) of an underlying set of semantic categories (B). Reading rules can be represented as reading a row in a matrix where the columns mark categories of grammatical meaning.

Agent	Motion	Location	Time
+S	+P	±Ad	±Ad

DIAGRAM I

This display can be interpreted, "Obligatory agent subject followed by obligatory motion predicate followed by optional location adjunct and optional time adjunct." Grammatical form and meaning can also be represented $+S_{ag} +P_{mo} \pm Ad_{loc} \pm Ad_{tm}$, if a one-dimensional display is more efficient.

There seem to be particular form-meaning associations, at least for particular languages. Thus, agent seems in English more closely related to subject than to predicate, motion seems more closely related to predicate than adjunct, and location seems more closely related to adjunct than either subject or predicate. We have seen, however, that categories can be variously represented in the surface structure, though complete interchangeability of form and meaning seems unlikely. Let us see how many ways we can represent the B string,

Agent   Motion   Location   Time

without changing aspect D (reference) from example to example.

Subject as agent: John runs in the park in the morning

Subject as motion: Running is what John does in the park in the morning.

Subject as location: The park is where John runs in the morning.

Subject as time: In the morning is when John runs in the park.

Notice that in English, putting into the subject tagmemes not usually found there requires a particular kind of sentence, one with the general form,

Subject + BE + Category Word + Clause

Indeed, the categories of grammatical meaning that are most easily established are those with almost unambiguous substitutes such as _who_ or _he_ (agent or goal), _where_ (location), or _when_ (time). Some other words in English can function like the substitutes in the pattern above:

The park is the _place_ John runs every morning.

or

Running is the _thing_ John does every morning.

We shall discuss this set of words in more detail below.

In Diagram I we labelled the columns with names for categories of grammatical meaning (aspect B) and filled the cells with names for grammatical forms (aspect A). The question now arises, what are the row vectors? Rows can, in light of the previous paragraph, be interpreted as different clause level syntagmeme types, distinguished both by grammatical form (presence or absence of object, presence or absence of adjunct in accompaniment column, etc.) and by the category in focus in the sentence. A category will be called "in focus" when it is subject in the surface structure. The matrix in Diagram I may now be expanded to include other focus-types involving the underlying pattern Agent Motion Location Time.

	Agent	Motion	State	Location	Time
Ag focus	+S	+P		±Ad	±Ad
Mo focus		+S/+C	+P		
Loc focus			+P	+S/+C	
Tm focus			+P		+S/+C

DIAGRAM II

There are two differences between this matrix and that in diagram I.

1) There is a new column label, marking a new category of grammatical meaning, *state*, referring here to the predicate type most usually represented by the verb *be*. That *state* is a different category from *motion* is indicated in part by the fact that they can not be conjoined:

*John runs and is in the park in the morning.

2) In the matrix in Diagram II, there are two symbols in some cells, indicating that a grammatical category may appear more than once in a reading. Thus in the second row, both subject and complement are readings of the category *Motion*, e.g.:

Running is what John does in the park every morning.

Here, Wh- +DO is a substitute for a motion verb, though it is ambiguous in that it can also substitute for some other categories of action (not State, however). The sentence above can be formulated as,

$$+S_{mo} \ +P_{st} \ +C_{mo}$$

Here as in our description of conjoining, the reference (aspect D) of the constituents with the same category marker is relevant. Both $+S_{mo}$ and $+C_{mo}$ have the same reference in this particular focus-type. This information must be made available somewhere in the grammar.

The reading of some rows in diagram II will result in an apparently misordered string, e.g.

$$+S_{mo}/+C_{mo} \ +P_{st}$$

The string may either be ordered properly by an obligatory permutation rule or the order above may prove of some benefit(for stating collocational restrictions) in that it groups tagmemes with common grammatical and lexical meaning.

In the reading with Mo focus, notice that the categories Agent, Location, and Time are not present in the surface structure of the clause,

$$+S_{mo}+P_{st}+C_{mo}$$

These categories have become embedded in the complement, which is a motion-focus dependent clause (i.e. "what John does in the park every morning"). These categories, then, will be recovered by the recursive reading rules when the generative process reaches the point where $+C_{mo}$ is given the exponent: Dep Mo-focus clause.

The simple matrix in diagram II will yield seven readings:

1) $S_{ag}P_{mo}$

2) $S_{ag}P_{mo}Ad_{loc}$

3) $S_{ag}P_{mo}Ad_{tm}$

4) $S_{ag}P_{mo}Ad_{loc}Ad_{tm}$

5) $S_{mo}C_{mo}P_{st}$

6) $P_{st}S_{loc}C_{loc}$

7) $P_{st}S_{tm}C_{tm}$

Readings 5,6, and 7 may, as has been mentioned, require obligatory permutation to the correct surface order. Readings 1 and 2 have only identity permutation, while Ad_{tm} may be optionally permuted to initial position in readings 3 and 4 and also before Ad_{loc} in 4. Each permutation may then be passed through optional conjoining rules.

The exponence rule for each reading will then require either 1) recursive back-looping, in case the exponent of a tagmeme is a syntagmeme of some or higher level, or 2) replacing each clause-level tagmeme with a phrase-level syntagmeme. Phrase level syntagmemes are read in the same way that clause level syntagmemes are, that is, in two-dimensional matrices. Many of the same categories of grammatical meaning recur from level to level, though they are manifested by different grammatical forms at each level. This is an essential point in the model we are now describing, though its importance may not appear until later. Thus, the category string Agent + Motion may be given different grammatical forms at clause, phrase, and word levels:

1) He runs ($S_{ag}P_{mo}$)

2) His running ($Sb_{ag}H_{mo}$) (Sb=subjunct, H = head)

3) Runner ($St_{mo}Sf_{ag}$) (St = stem, Sf = suffix)

Grammatical forms are distinctive for each level, and levels may now be defined by the grammatical form that a given category of grammatical meaning takes. Also, each level seems to have certain distinctive categories that do not appear at other levels. Thus, for sentence level the slots *concessive* and *conditional* seem distinctive. That is, neither the concessive category (as in "Although water is now plentiful, some take sponge baths" where "although water is now plentiful" is a clause representing the category *concessive* in a sentence syntagmeme) nor the conditional category (as in "If he comes, count me out") are represented at other levels of the grammar. Various categories seem distinctive to the phrase level; that is, there does not seem to be a clause level parallel to the functions of "the" and "a". On the clause level, constructions with the category *state* (forms of copula *BE*) seem distinctive. It may turn out, however, that all categories are potentially present at all levels and that the only distinction between levels is the form, not the meaning of the syntagmemes.

For the sake of illustration, let us assume a simple structure for a phrase level syntagmeme which might be an exponent of S_{ag}:

Agent-focus phrase + $+\mathrm{Det}_{spec} \pm \mathrm{Sb}_{qual} + \mathrm{H}_{ag} \pm \mathrm{R}_{loc}$

(E.g. The old man near the door.)

We can present this phrase-level reading as the reading of a row in a phrase-level syntagmeme matrix:

	Specification	Quality	Agent	Location
Ag focus	+Det	± Sb	+H	±R

DIAGRAM III

Though we will not go into detail here, a phrase--like a clause--may have location focus (Those near the door, $+Det_{spec}$ $+H_{loc}$) or quality focus (The old, $+Det_{spec}$ $+H_{qual}$), or other types of focus. Permutations are both obligatory and optional at this level, too, and affect the exponence rules in interesting ways; e.g. $+H_{ag}$ $+Sb_{loc}$ can have as an exponent "man near the door"; $+R_{loc}$ $+H_{ag}$ can have only single words as exponents: "nearby man," or "nearest man", for instance.

After a clause level tagmeme had been replaced by a back- or forward-looping exponent, the output of the grammar (intermediate string) might take this form (for the simple string we have been using to illustrate the model):

$$S_{ag}(Det_{spec}\ Sb_{qual}H_{ag}R_{loc})\quad P_{mo}(.......)\ Ad_{loc}\ (.....)\ Ad_{tm}(.....)$$

That is, each reading operation describes the internal structure of a previously read constituent of a syntagmeme. The final output of the grammatical section of the grammar will be a string of morphemes and morpheme classes N/V/Adj. carefully marked for a cross-reference lexicon. For example, using the formula above (which has not yet undergone word level operations), the lexical item which is head of the subject phrase will carry into the lexical component the markers +Ag +Spec +Qual +Loc +Mo: that is, the categories of grammatical meaning represent the grammatical constraints on lexical collocation. If properly formulated, the categories should provide a very delicate cross-reference system in which the item "man", for instance, might be listed as

N/Ag (+Qual ____ +Loc +Mo)

(Man = noun, agent, in context: quality preceding, location and motion following)

Of course, "man" may also be goal or action, and so the above would not be the only entry for the lexical item, "man".

In this brief sketch of the form of a grammar, many important questions have been left out. Hopefully, many of these will be answered by the fragment of English grammar that follows. Some will remain unclear. For instance, it is not clear just how information about aspect D will be added to the grammar. We can easily add markers to the reading rules that signal "select different lexical form with same reference" to yield a sentence like,

John is my brother.

However, what does this signal mean when we have not yet specified in detail a structure for the lexicon? What might be called "equivalence of reference" remains to be defined in any operational way.

Another important question will remain unanswered here, though a tentative answer can be suggested. Are forms like questions or imperatives to be generated by reading rules or by something like transformations? It seems simpler at present that these forms be generated by the reading rules, which raises the possibility of a single surface form representing two underlying categories, just as we have seen examples of two surface forms representing a single underlying category (e.g. The park is where John runs every morning). Hence, in a reading for a question we may have two readings of, for example, subject: $\text{Subject}_{\text{agent}}/\text{Subject}_{\text{ques.}}$ and a later exponence rule that reads this as John/who...... with surface deletion of John. It is not clear, however, whether question is a deep or a surface phenomenon. Likewise, imperative forms may merely involve a surface deletion of an underlying subject tagmeme.

2.4 Summary

We have seen that Longacre's model, though it has obvious weaknesses, is a model for a generative grammar based on string constituent analysis of constructions. The weaknesses of Longacre's model are chiefly the result of an incomplete description of string constituents (i.e. tagmemes). That Longacre's notion of the tagmeme is inadequate (marking only aspects A and C) is demonstrable if we try to describe conjoining, even of a very simple sort. We have here described an expanded version of Longacre's model which restores aspect B to the tagmeme and suggests adding aspect D in some as yet unclear way. Reading rules have been shown to select surface forms for underlying combinations of categories of grammatical meaning. These categories were seen to be represented by different forms at different levels of the grammar, thereby providing a clear separation of grammatical levels. The categories and surface forms are finally given fillers (or exponents, in Longacre's terms) which are particular lexical items or labels for lexical classes (aspect C). It was suggested that these lexical classes could be subclassified by using category labels as context markers for particular lexical items, thus providing complex grammatical constraints on lexical collocation.

CHAPTER III

ENGLISH SUBJECT TAGMEMES: CLAUSE LEVEL READING RULES

In this chapter we will first list, illustrate, and comment on the English subject tagmemes and then demonstrate how these tagmemes might be generated in clause level reading rules. In the following chapter we will describe the exponents of these tagmemes and the phrase level readings of some of them. It is not assumed that this list of subject tagmemes is complete or that the reading rules are exhaustive; however we have tried to list some of those apparent subject types which this grammar cannot at present accommodate. Furthermore, this approach to the English subject tagmeme is only incidently and rarely either comparative or historical, even though we feel strongly that the categories of these tagmemes represent real language universals and have deep historical roots.[1]

3.1 English Subject Tagmemes[2]

The English subject tagmemes are described here in terms only of aspects A and B of the complex tagmeme. The relevance of aspect D to the readings rules will be discussed in 3.2.9 below. Aspect C is supplied by terminal exponenee rules as labels for lexical form classes and will be discussed in Chapter IV.

3.1.1 Subject as agent (S_{ag})

(1) John runs in the park every morning.

(2) Whoever was in the room killed Bill.

(3) One of them came.

3.1.2 Subject as proposition (S_{prop})[3]

(4) The idea that John is frightened amuses me.

(5) The idea amuses me.

(6) That John is frightened never occurred to me.

(7) It never occurred to me that John is frightened.

(8) There is no doubt that John is frightened.

3.1.3 Subject as instrument (S_{in})

(9) A stone struck Bill.

(10) The key opened the door.

(11) His left hand brushed away the crumbs.

(12) His superior intellect won the game.

3.1.4 Subject as goal (S_{gl})[4]

(13) The book burned.

(14) The book was burned by John.

(15) Bill was hit by John.

(16) The tree grew.

(17) The grass smells fresh.

(18) The typewriter works fine.

3.1.5 Subject as time (S_{tm})

(19) 1957 was a vintage year.

(20) The morning is the best time for swimming.

(21) Sunday passed quickly.

3.1.6 Subject as location (S_{loc})

(22) From here to there is about ten miles.

(23) Ann Arbor is a nice place to live.

(24) The south side is warmer.

(25) It is warmer on the south side.

(26) There are many houses on the south side.

3.1.7 Subject as quality (S_{qual})

(27) Kindness killed the cat.

(28) Sincerity may frighten the boy.

(29) His eagerness to please is embarrassing.

(30) His anxiety for news overcame his timidity.

(31) The possibility of finishing seems remote.

3.1.8 Subject as nexus (S_{nex})[5]

(32) For John to be frightened is ridiculous.

(33) It is ridiculous for John to be frightened.

(34) Even to begin is to fail.

3.1.9 Subject as state (S_{st})

(35) Being cold is no fun.

(36) His being there made no difference.

(37) Our being married at home bothered my mother.

3.1.10 Subject as act (S_{act})[6]

(38) Playing tennis is fun.

(39) Constructing the building took all summer.

(40) The construction of the building took all summer.

(41) John's love of sports is obsessive.

(42) Your plan to start in September failed to get support.

3.1.11 Subject as motion (S_{mo})

(43) Walking is healthy.

(44) Dancing with Mary is a privilege.

(45) Running five miles a day strengthens the heart.

3.2 Discussion of set of English subject tagmemes

3.2.1 Discovery procedures and criteria for establishing categories

Though Chomsky and Lees are entirely right that there are no automatic discovery procedures, infallible recipes for analysis, or taxonomic definitions for identifying three-legged zebras with no stripes,[7] there is some need to insure that the units, constructions and rules one posits in a grammar are not ad hoc but have what Pike has called emic relevance for the language. The trouble with a grammar made up of ad hoc categories and rules is not that it can't generate language but that it gives us no real insight into the workings of a particular language or languages in general -- which is, after all, the purpose of linguistics, for we already have the competence to generate language with much more complexity than any grammar can ever hope to match. In order to make it more likely (not assure) that tagmemes and syntagmemes are not ad hoc, both Pike (1962) and Longacre (1964a, pp. 17-23) have presented procedures for distinguishing emic units from etic variants. In general, both require that two etic units differ in at least two significant ways before they may be established as emically contrastive. Pike and Longacre differ over just what a significant contrast is. Both agree that a significant difference might be linear ordering, different tagmemes in a syntagmeme (especially nuclear tagmemes), different exponents, different transform potential, or different history of derivation in the generative process. Pike also considers that distributional differences between tagmemes and syntagmemes are significant in establishing emic units.

Each of the English subject tagmemes we have established differs from the others in exponent (filler type) and in category manifested. Because categories like agent, instrument, goal, etc. are often covert in modern English (though there were overt markers for some of them in Old English) it is difficult to give precise criteria for establishing them except that they be relevant to more than one sort of rule in the grammar. Thus, most of the categories have been posited for English when readings of a category (e.g. S_{ag} or Ad_{in}) are not conjoinable with formally equivalent readings of another category (e.g. S_{gl} or Ad_{loc}). For example, as we saw in Chapter II, S_{ag} and S_{in} are not conjoinable, even when in identical contexts:

John hit Bill

The stone hit Bill

*John and the stone hit Bill

On this evidence, and on the evidence of differing exponents, we establish S_{ag} and S_{in} as emically different tagmemes. For a second example, S_{ag} and S_{gl} are clearly contrastive on the evidence of the following set of sentences:

John cooked.

The potatoes cooked.

*John and the potatoes cooked.

This criterion of non-equivalence in reference to conjoining was effective in distinguishing almost all of the tagmemes above.

However, some contrasts have to be established obliquely or cannot be clearly established and remain tentative. It is important to identify these. In some cases tagmemes of the category proposition (see 3.1.2 above) can be conjoined with tagmemes of different categories so as to yield clearly acceptable sentences:

> Bill accepted the trophy and the recognition that he was the best tennis player in the state. (O_{gl} and O_{prop})
>
> Sincerity and the assumption that another may be right are admirable qualities. (S_{qual} and S_{prop})

In spite of this apparent overlap with two other categories in conjoining, the category prop seems strongly distinctive in its exponents and in its permutation potentiality (e.g. examples 7 and 8).

On the other hand, the categories whose exponents are verbs or nominalized verbs (state, act, motion) can be assumed only very tentatively. It is as yet unclear how many categories underly the English predicate[8] and hence there is no assurance that state, act, and motion will turn out to be deeply relevant in the grammar of the predicate. We have not attempted to analyze the predicate deeply here and establish these categories only because they seem to be relevant to the subject. They often cannot be conjoined at clause level in the predicate, as in the following examples:

> *He is and plays there. (Pr_{state} and Pr_{act})
>
> *He is and walks in the meadow (Pr_{st} and Pr_{mo})
>
> *He likes and walks Sundays (Pr_{act} and Pr_{mo})

But they usually can be conjoined in the subject:

> Playing tennis and dancing are my favorite forms of exercise.
>
> John's being there and his avoidance of Mary were the two most important factors in his later conviction.

These categories can only be tentatively established and the description of these subject tagmemes will have to be revised when the predicate has been more adequately described.

Another important way of contrasting categories is in terms of their substitutes and what we called in Chapter II their "category words." Thus the question substitute for in is "with what," for accomp "with whom," for tm "when," for loc "where," etc. Some categories (e.g. in) cannot be replaced by he, she, or who, while others are neutral in regard to the animate-inanimate distinction (e.g. gl). Moreover, the category labels themselves and similar words seem to play a special role in language, especially in the pattern which we can roughly describe as Subject + Pr_{state} + Category Word + Relative Clause:

> The park is the place where I run every day.
>
> Morning is the time when I feel the best.
>
> A pencil was the instrument with which he opened the door.
>
> Kindness is the quality that he admires the most.
>
> Being cold is the state that bothers him most.
>
> Etc.

Notice that in all these sentences Subject, Category word, and Relative substitute have the same reference (aspect D), a point we shall discuss later (3.2.9).

These substitution possibilities are of great importance in discourse analysis, as is the description of anaphoric relationships.[9] They will not be of special concern in the remainder of this grammar, however.

3.2.2 Categories and Lexical Subclasses

One of the major problems in writing a grammar is to decide how much semantic information belongs in the grammar and how much should properly be provided by the lexicon. Few would argue that restrictions do not belong in the grammar to prevent such collocations as

*The rice are gentlemen.

In order to describe these collocational constraints most grammarians (e.g. Jespersen, Fries, Nida, Chomsky, Lees, Pike, and Longacre) define subclasses of lexical classes which are marked $\text{Noun}_{\text{mass}}$, $\text{Noun}_{\text{animate}}$, $\text{Noun}_{\text{human}}$, $\text{Verb}_{\text{transitive}}$, etc. As we observed in discussing generative-transformational grammar, these subclassifications reflect the reference (aspect D) of grammatical constituents. That is, terms like animate, mass, human, etc. refer to the world outside language, or rather, people's conceptions of this world. It seems reasonable to us that these reflections of our conceptualization of the world around us be a part of the deepest stratum of a grammar, not a part of the surface structure, and that they be marked early in the grammar. The fact that in English inanimate objects don't normally walk, and talk, and believe things is a part of our universe of discourse. In fairy tales and fantasy these things happen and our universe of discourse changes. Thus, we would explain the following sentence at a deep level of the grammar.

The moon wept to see such misery.

Rather than say that a lexical item of the subclass marked inanimate is here in a context that requires an animate subject (and hence that the sentence is deviant), we would say that the sentence has a deep structure

$$S_{loc}{}^{Pr}{}_{act}{}^{Ad}{}_{reason}$$

The deviance of the sentence is thus a feature of the deep structure where $S_{loc}{}^{Pr}{}_{act}$ reflects a special universe of discourse. In this approach the deviance is a property of the underlying string, not of the lexical units, a property of the meaning (aspects B and D) not of the form (aspects A and C).

Hence, and in so far as possible, we shall try to avoid semantic subclassifications of lexical form classes (exponents) and try to put these collocational restraints into the reading rules. Notice that the categories we have established carry inherently a great deal of the information formerly put into the labels of lexical subclasses. For example, the category agent implies the inherent semantic features animate and concrete, the category instrument implies the feature inanimate but is neutral in regard to the distinction between concrete and abstract (e.g. example 12), the category goal is neutral re animate-inanimate but implies the feature concrete, the category proposition implies the feature abstract, and the category motion implies the feature intransitive.

We will not specify here all the features that are inherent in each category; it is probably impossible at this point. It is our hope that semantic constraints on collocations of lexical items can be stated entirely in terms of the underlying structure (see section 4.6 below).

3.2.3 Rank in a Nexus

We have deliberately used Jespersen's language in describing this feature of some of the English subject tagmemes, for it involves a justification of one of his insights which has been rejected by some linguists.[10] As we pointed out earlier, Jespersen commented on the agnate relation of the nexus,

The dogs bark.

and the junction,

The barking dogs. . .

In assigning rank to the constituents of the junction (barking= secondary, dogs=primary) he carried over this ranking to the constituents of the nexus, ranking the subject primary and the predicate secondary.

Example 2 clearly illustrates a nexus in which the subject of the embedded nexus is primary:

(2) Whoever was in the room killed Bill.

The substitute for the subject, an embedded clause, is he or she, as in the sequence,

Whoever was in the room killed Bill. He must have.

Unless the subject is reduced from a primary, embedded clauses are in this sense headed constructions. That is, they are replaceable by a substitute for the subject. There are many means to reduce a subject. Some of them are (a) to put the subject into the so-called "possessive" form, as in

(36) His being there made no difference.

(41) John's love of sports is obsessive.

(b) to put the subject into something analogous to a "dative" form, as in

(32) For John to be frightened is ridiculous.

or (c) to prepose to the subject a marker which puts focus either on the whole nexus (one function of that) or on some tagmeme other than subject. The following illustrate these two kinds of subject reducing markers:

(6) That John is frightened never occurred to me.
Wherever John goes makes no difference.
Whenever he goes makes no difference.

Notice that if the subject of the embedded clause is reduced, the substitute for the subject of the embedding clause is it, this, or one of the set of prop substitutes: idea, fact, statement, notion, feeling, hope, etc. (See 3.1.2) And if it is reduced by means (a) or (b), it becomes an optional constituent of the embedded clause and clearly descends in rank (in Jespersen's sense) from a primary to a secondary or a tertiary. Jespersen's insight was correct: there is rank in a nexus.

3.2.4 Rank reduction and obligatorily excluded tagmemes

Just as a subject may be reduced in rank in an embedded clause, some tagmemes are mutually exclusive and hence one of them is obligatorily reduced in rank if they both are to occur in a string. For example, in a clause with S_{in}, the category agent may not occur as a clause level constituent. Thus, if we have the sentence,

(9) A stone struck Bill.

there is no way to include agent as a clause level constituent. That is, there is no pair of sentences in English like,

John struck Bill with a stone

*A stone struck Bill by John.

Agent must be introduced at a lower level and, hence, descend in rank, as in,

A stone thrown by John struck Bill.

Or, agent may be introduced in another clause:

John threw a stone, and it struck Bill.

Notice, however, that S_{ag} and the category instrument are not exclusive. Another example of this phenomenon excludes the category prop following S_{in}.[11]

3.2.5 Passives

In this grammar passives are generated by reading rules, not by transformations of surface structure. That is, we do not derive passives from underlying active syntagmemes. The relation between them lies not in surface manipulations but in the underlying deep structure. Clearly not all objects can become subjects of transitive verbs. Since the passive transformation cannot be described in terms of surface labels like subject, object (aspect A) or NP, VP (aspect C), constraints on the passive transform have usually been stated in terms of semantically marked lexical subclasses. Consider how many subclasses would be necessary to prevent the following strings:

> *His shins were bruised by John. (Where his refers to John)
>
> *A book is had by John.
>
> *Bill is resembled by John.
>
> *Shakespeare was learned about by John.

Our method of describing passives directly in the reading rules avoids the problem of defining complex, ad hoc sub-classification of lexical filler classes (See 3.2.2) or, alternatively of describing complex context restrictions in the reading rules.

3.2.6 Dummy subjects

Once tagmemic rules for sentences and paragraphs are written, it is likely that permutation rules (and some reading rules) can be made context sensitive. That is, there are contextual reasons, some of which we can describe, for most permutations such as

> Yesterday I bought a new suit.
>
> Of all the singers, John was the best.

It has been observed that in English discourse the first part of the sentence (usually the subject) relates the sentence to the previous context, while the second part presents new information. Thus in a discourse the function of subject is frequently to maintain focus on a topic. While one sort of focus (handled by reading rules) brings the topic in focus to the subject position, another kind of focus or defocus (described by permutation rules), moves a tagmeme out of subject position and replaces it by a dummy subject, either it or there. In such permutations we can say that the dummy subject anticipates a postponed subject or that it has the same reference as the subject, but that the subject has moved out of focus (keeping, however, its rank). Examples (6) and (7) illustrate this defocusing phenomenon:

(6) That John is frightened never occurred to me.

(7) It never occurred to me that John is frightened.

A strong formal reason for thinking that it in (7) refers to the postponed clause is that, as we saw in 3.2.1, tagmemes can have as anaphonic substitutes pronouns and category words; however, a tagmeme cannot be replaced anaphorically by both a pronoun and a category word at once, unless some such device as parenthesis is used:

That John is frightened never occurred to me
It (the idea) never entered my mind.

We have seen, however, that a category word and a that-clause can occur together:

(4) The idea that John is frightened amuses me.

But we cannot have the construction,

*It amuses me the idea that John is frightened.

which suggests that it has replaced the category word idea, an interpretation which is strengthened by the occurrence of another, less common permutation,

The idea amuses me that John is frightened.

The hope lingers that he will return.

We would give a similar interpretation to example (25):

(25) It is warmer on the south side.

That is, we interpret (25) as a defocussing permutation of

(24) The south side is warmer.

In both cases the subject is S_{loc}.

The interpretation of the dummy subject there is much more complex. The complexity can be suggested if we examine the following set of sentences:

(a) In the room are four windows.

(b) *Four windows are in the room.

(c) There are four windows in the room.

Now let us add tag questions to each, a traditional procedure for identifying the subject:

(a') In the room are four windows, aren't there?

(b') *Four windows are in the room, aren't they?

(c') There are four windows in the room, aren't there?

Because the subject of the tag questions in (a') and (c') is there in both cases, we might interpret there as a substitute for S_{loc}.

However, in each case it is "four windows" which governs the number of the verb (another traditional test of subject). Hence we might say that in (a) and (c) and (a') and (c') the subject is "four windows." However, if "four windows" is the subject, why don't (b) and (b') occur? The problem is identical to the one raised by the pair,

(8) There is no doubt that John is frightened.

(8') There are some doubts that John is frightened, aren't there?

* Some doubts are that John is frightened.

For the present it seems most reasonable to consider *there* a dummy subject for S_{prop} and S_{loc}, and add a rule at word level that in the context #there______, BE is governed by the following tagmeme. This is an unsatisfying, ad hoc rule. It works but it gives us no great insight into the workings of *there* as a dummy subject.

3.2.7 *Anaphoric deletions*.

Most, if not all, deletions can be handled as markers of anaphoric relations between clauses. Though many grammarians have discussed ellipsis, only two (to my knowledge) have pointed out this anaphoric use of deletion. In his article, "Signe zero" Jacobson writes,

> Nous sommes tentes d'interpreter l'ellipse plutôt comme une sous-entente des termes anaphoriques qui "representent" le contexte ou bien des termes deictiques qui "presentent" la situation...L'ellipse est donc signe anaphorique ou deictique zero.

The suggestion in Jacobson's remark is developed in detail by Karlsen in Studies in the Connection of Clauses in Current English. Their suggestion is adopted here to explain such subjects as the following:

John's is on the table.

A few arrived.

Those I've finished are on the table.

In each case the context would supply us with the missing head word, which can be said to be present as 0 and reconstructable, in part, from the rest of the sentence. Thus in each of the examples above, the deleted subject is concrete (thus excluding S_{prop}, S_{tm}, S_{qual}, S_{nex}, S_{st}, S_{act}, and S_{mo}) and in the second example (A few arrived) it is also animate, narrowing the selection down to S_{ag}. (This argument presupposes the presentation of the reading rules in Section 3.3 below.)

However, anaphoric deletion will not explain subjects like the following, which are always plural and which do not seem to have the same context restrictions as the examples above.

The good die young.

The brave deserve the fair.

We shall consider (4.3.1) such adjective-headed subject phrases as one exponent of S_{ag}.

3.2.8 Linear and Nesting Recursiveness

There are two kinds of recursiveness in a generative grammar which reads from larger to smaller units (e.g. from clause to word). The first kind, which we called linear recursiveness, refers to various forms of conjoining and always adds to the number of tagmemes in a syntagmeme. Linear recursiveness is the repeating of a tagmeme; for example, the subject might be repeated:

$S_x Pr_y O_z$ $\quad$ S_x and $S_x Pr_y O_z$

Nesting recursiveness, on the other hand, refers to the reading of a syntagmeme at any level as a component of a same or lower level syntagmeme; that is, a clause may be a component of a clause or a clause may be a component of a phrase. Example (2), for instance, illustrates a clause embedded within a clause:

(2) Whoever was in the room killed Bill.

Here S_{ag} has as its exponent the clause,

Whoever was in the room

We saw in 2.2.3 above that the subject of the embedded clause was primary (as it is in all clauses unless reduced) and that we might call it a subject-focused clause; its replacement is he/she, pronouns which replace subjects of clauses.

Example (42) also illustrates nesting recursiveness, this time with a clause in which the predicate is primary, what we might now call a predicate-focused clause:

(42) Your plan to start in September failed to get support.

The subject here is S_{act} and its exponent is a clause in which subject has been reduced to make it predicate-focused:

Your plan to start in September

This predicate-focused clause is agnate to the subject-focused clause:

You plan to start in September.

This clause has the structure:

$S_{ag}Pr_{act}O_{nexus}$

This reading might also yield,

You plan for Bill to start the book in September.

That is, O_{nexus} may include a predicate-focused (to + V) clause with a subject, object, and various adjuncts. In the predicate-focused clause,

to start in September

only Pr_{act} and Ad_{tm} have been read. What is the subject in (42) of the predicate "to start"? It is somewhat ambiguous, but most likely "you." The ambiguity is apparent if we consider that example (42) might be spoken by an employee to an employer, where it might mean either,

Your plan (for you) to start in September failed
to get support

Or,

Your plan (for us) to start in September failed
to get support.

Notice that the absence of the subject of the nexus (for you/for us) is in a sense anaphoric; its absence automatically signals that the subject is either the nearest agent to the left (most likely option) or clearly understood from the previous context.

Failure to give a correct structural description to nested syntagmemes has been a serious charge against tagmemic grammars. As Postal (1966, p. 95) has recently stated it,

> Categorizing some NP as subject characterizes the subject relation uniquely only if there is a single predicate present. Furthermore, one must recognize that position in the actual sentence in no sense uniquely indicates what bears what relations to what. For instance:
> (5) John is anxious to criticize
> (6) John is difficult to criticize, it is difficult to criticize John, To criticize John is difficult
> In (5) John is understood as subject of both anxious and criticize. But in all the examples of (6), despite the superficial identity of the structure of the first with (5), John is the object of criticize....

Postal's criticism is a profound one and reflects the frequent failure of tagmemic grammars to distinguish clearly between surface and deep structure.[12] Even given Longacre's model, however, Postal fails to understand the distinction between reading and exponence rules or to consider that his (6) includes a permutation. That is, the subject of (6) is "to criticize John" S_{nex} and the subject of (5) is "John" S_{ag}. In our analysis, the proper structural description of (5) is,

$$S_{ag}Pr_{st}C_{qual} \text{ where } C_{qual} = \text{"anxious to criticize"}$$

$$(H_{qual}+Sb_{nex})$$

The proper analysis of (6) is,

$$S_{nex}Pr_{st}C_{qual}$$

To make Postal's examples parallel in complement, we would rewrite them,

P(5) John is (anxious to criticize)

P(6) To criticize John is (difficult to do)

Hence, in (6) John is not the subject in the deep structure, but, rather, the object of a nexus (i.e. a verb-headed clause). Furthermore, the subject of "to criticize" in (5) is ambiguous, though _most likely_ the nearest agent to the left ("John"). That is, the most likely reading is,

John is anxious (for himself) to criticize (someone).

but a possible reading is,

John is anxious (for us) to criticize (someone),
where "us" is inclusive.

Hence Postal's (5) is similar to our (42) discussed above. This ambiguity is clearer, perhaps, in,

John is anxious to play.

Here only context could tell us whether the subject of "to play" is John as it might be in,

John is anxious (for himself) to play (the guitar)

or _we_ (inclusive),

John is anxious (for us) to play (basketball).

Our reformulation of Longacre's model allows us to answer the repeated charge that tagmemics only labels surface structure. A tagmemic grammar can be adequate to the description of both linear and nesting recursiveness.

3.2.9 Some Remaining Problems

Before presenting reading rules for the English subject tagmemes listed above, we will discuss two subject tagmemes whose analysis remains problematic. First, it is not entirely clear how we would describe the subject of the sentence,

(46) The book impressed me.

There are many sentences of this sort in which an inanimate, concrete subject seems almost an agent. In some cases we can analyze such a subject as S_{in} as in,

(47) The book hit the window

which is agnate to,

(48) Someone hit the window with a book.

In other cases, such subjects are clearly S_{gl}:

(49) The book looks interesting (to me).

Here the agent, the "one who looks," is present as an adjunct (Ad_{ag} = to me), and the sentence is probably agnate to,

(50) I look at the book with interest.

It is possible to interpret the subject of (46) as an instance of S_{in}, for it is not possible to express agent in (46) which, as we saw in 3.2.4 above, is a feature of clauses with S_{in}. In this case we would consider the agnate agent-focused clause as,

(51) I impressed myself with the book.

Then, the subject of the sentence,

(52) The book costs (me) five dollars.

would also be an instance of S_{in}. The agent-focused agnate to (52) then would be,

(53) I cost myself five dollars with the book.

This seems quite awkward and, hence, unlikely, though no better solution has appeared yet. Furthermore, to consider all instances of inanimate, concrete subjects as S_{in}, S_{gl}, or possibly S_{loc} does not explain the subject of a sentence like,

(54) The book fell on the floor.

Perhaps another subject tagmeme remains to be identified.

The second subject tagmeme whose identification must remain problematic is that illustrated by the sentence,

(55) John is a good man.

The surface structure of this sentence is Subject + Predicate + Complement. In Diagram II (Chapter II, page 71) we analyzed such patterns as including two appearances of a category in a reading. For example, the sentence,

(56) The park is the place where John runs every morning.

was assigned the structure,

$$S_{loc}Pr_{st}C_{loc}$$

Hence it is possible to analyze (55) as

$$S_{ag}Pr_{st}C_{ag}$$

It may seem puzzling that we can consider the subject of a stative verb to be an agent. There are several reasons why we have chosen this interpretation rather than the alternative of recognizing a different subject type, say, following Halliday (1967), Subject as attribuant. There are four reasons for rejecting this alternative:

1) Internal structure. Any of the tagmeme types we have established may occur as subject of a stative verb. The suggested tagmeme S_{attrib} would in this sense be the sum of all possible subject types but have no distinctive features of its own.

2) Undifferentiated by the criteria. There are no substitution potentialities or constraints on conjoining by which to contrast the category attribuant from any of the other categories. That is, there is no "category word" for attribuant as there are for the other categories (e.g. who, what, when, where, the time, the place, etc.). Establishing S_{attrib} would greatly complicate conjoining rules, for as S_{attrib} is the sum of all subject types, then any two subject types could be conjoined before Pr_{st}, which is clearly not the case:

> *John and that he might be frightened are important considerations.
>
> *Yesterday and here are important considerations.

As a matter of fact, conjoining constraints before Pr_{st} can be described as in all other clause readings -- in terms of the categories we have already established.

3) Contrastive features not in subject. The difference between the sentences,

(55) John is a good man

and

(57) John saw a good man

is not in the subject types but in the predicate and what follows the predicate. In (55) the predicate is Pr_{st} followed by a complement, which, in our analysis, must be read as the same category as the subject, in order to prevent,

*John is the place I like best.

*The idea is a good man.

In (57) the predicate is Pr_{act} followed by an object which has no category relationship with the subject. We can explain the difference between these sentences without positing different subject types.

4) Analogy. If example (56) has S_{loc} rather than S_{attrib} it seems consistent to read (55) as S_{ag}. Perhaps the difficulty is only in the term agent. If so, we would prefer to have the reader bear with our broadening of the term (there seems no better one) rather than resort to purely abstract labels for the categories (e.g. numbers or symbols of some sort).

There is an interesting parallel here with traditional formal logic where all statements must be given the grammatical form subject-copula-complement. Without getting into the thorny problem of concrete universals and their interpretation in modern symbolic logic, we would, in traditional logic, interpret the statement,

John killed the tiger

as something like

John is the man who killed the tiger.

In traditional logic these statements are seen as equivalent. It must be stressed that we are not here attempting to justify our analysis by using the rules of traditional logic; our justification is given in the four points listed above. What we are attempting to show is that our labelling of the subject of a stative verb as agent is not entirely without precedent.

It should also be noted that our category agent is in no way equivalent to Fries's (1952) subject as "performer." For one thing, Fries was not establishing grammatical categories as separate from grammatical forms, and he described his subject types in terms of the particular verbs with which they could collate in surface level substitution tests. Fries, for instance, described (p. 178) "the chair" as "performer" in

The chair tipped over.

In our analysis the subject here is S_{gl}. We are not labelling lexical distribution classes in terms of surface substitution in frames, though this technique is a very useful heuristic in some cases, but rather we are attempting, in part, to derive grammatical specifications for these classes via generative rules, so that, in a sense, the grammar selects sets of lexical items as realizations of terminal grammatical strings. This difference has been discussed above (section 3.2.2) and will be illustrated briefly later (section 4.6).

However, aspect D (lexical meaning) does play a part in the grammatical description of a reading like

$S_{ag}Pr_{st}C_{ag}$.

As Jespersen (1965) pointed out, in a sentence like

(58) A good man is John.

it is part of our linguistic competence to recognize that this is an inversion, that "John" is still the subject. As Jespersen stated (p. 150),

> If one of the substantives is perfectly definite, and the other not, the former is the subject; this is the case with a proper name.

Furthermore, if both subject and complement are equally definite, then either may be subject:

(59) Eight o'clock in the morning is the best time to go skating.

(60) The best time to go skating is eight o'clock in the morning.

Hence, the grammar must specify the exponence of subject and complement in this pattern so that it is true that x (subject) is a y (complement) or x (subject) has identical reference with (=) y (complement). To do this is to mark aspect D (lexical meaning or reference) in the reading rules. The only way we know how to do this at present is to mark tagmemes in the reading rules, when necessary, with superscripts:

$$S^{x}_{ag}\ Pr_{st}\ C^{y}_{ag}$$

This will carry into the lexical component of the complete grammar the instructions: select lexical items if and only if x is a y or x=y. This is an incomplete but well-motivated solution; at least the superscripts will keep before us the reminder that one aspect of the tagmeme has not been adequately described and that there is a hole in our model waiting to be filled.

3.3 Reading Rules for English Subject Tagmemes

We have not yet identified all the categories that may occur in an English clause. Some (e.g. manner, accompaniment) will be introduced below. Others (e.g. reason, source) will not be discussed here but rather must await a fuller treatment of the English clause before they are explicated. We are interested here only in rules sufficient to generate the English subject tagmemes. Hence the reading rules below are merely illustrative, sufficient only to provide readings for the syntagmemics and syntagmemes underlying examples (1) through (45) above.

In the matrix below (Diagram IV) the columns mark categories of grammatical meaning (aspect B) of the tagmemes. The rows distinguish the eleven subject tagmeme types, though here each type will be represented in several rows corresponding to the illustrations form (1) to (45). In the cells are symbols for grammatical form (aspect A) These are S 'subject,' Pr 'predicate,' O 'object," C 'complement,' and Ad 'adjunct.' Optional tagmemes selected for an illustration will retain the ± symbol to show that they were options. Options not selected for a particular example will be shown in parentheses.

Reading Rules for English Subject Tagmemes

Diagram IV

Categories of Grammatical Meaning

		Dummy (there) 1	Ag 2	Prop 3	Nex 4	Mo 5	Act 6	St 7	Gl 8	In 9	Qual 10	Loc 11	Tm 12	Man 13	Accomp 14
S_{ag}	(1)		S			Pr			(±0)			±Ad	±Ad	(±Ad)	(±Ad)
	(2)		S				Pr		0	(±Ad)		(±Ad)	(±Ad)	(±Ad)	
	(3)		S			Pr						(±Ad)	(±Ad)	(±Ad)	(±Ad)
S_{prop}	(4)			S			Pr		0			(±Ad)	(±Ad)	(±Ad)	
	(5)			S			Pr		0			(±Ad)	(±Ad)	(±Ad)	
	(6)			S			Pr		0			(±Ad)	(±Ad)		
	(7)	PERMUTATION OF (6)													
	(8)	X		S				Pr							
S_{in}	(9)						Pr		0	S		(±Ad)	(±Ad)	(±Ad)	
	(10)						Pr		0	S		(±Ad)	(±Ad)	(±Ad)	
	(11)						Pr		0	S		±Ad	(±Ad)	(±Ad)	
	(12)						Pr		0	S		(±Ad)	(±Ad)	(±Ad)	
S_{gl}	(13)						Pr		S			(±Ad)	(±Ad)	(±Ad)	
	(14)		±Ad				Pr		S	(±Ad)		(±Ad)	(±Ad)	(±Ad)	
	(15)		±Ad				Pr		S	(±Ad)		(±Ad)	(±Ad)	(±Ad)	
	(16)						Pr		S			(±Ad)	(±Ad)	(±Ad)	

		Dummy (there) 1	Ag 2	Prop 3	Nex 4	Mo 5	Act 6	St 7	Gl 8	In 9	Qual 10	Loc 11	Tm 12	Man 13	Accomp 14
	(17)		(±Ad)				Pr		S		C	(±Ad)	(±Ad)		
	(18)		(±Ad)				Pr		S			(±Ad)	(±Ad)	±Ad	
S_{tm}	(19)							Pr				(±Ad)	S^x/C^y		
	(20)							Pr				(±Ad)	S^x/C^y		
	(21)						Pr					(±Ad)	S	±Ad	
S_{loc}	(22)							Pr				S^x/C^y	(±Ad)		
	(23)							Pr				S^x/C^y	(±Ad)		
	(24)							Pr			C	S	(±Ad)		
	(25)	PERMUTATION OF (24)													
	(26)	X						Pr	C			S	(±Ad)		
S_{qual}	(27)						Pr		O		S	(±Ad)	(±Ad)	(±Ad)	
	(28)						Pr		O		S	(±Ad)	(±Ad)	(±Ad)	
	(29)							Pr			S/C	(±Ad)	(±Ad)		
	(30)						Pr				S/O	(±Ad)	(±Ad)	(±Ad)	
	(31)		(±Ad)				Pr				S/C	(±Ad)	(±Ad)		

Diagram IV (continued)

		Dummy (there) 1	Ag 2	Prop 3	Nex 4	Mo 5	Act 6	St 7	Gl 8	In 9	Qual 10	Loc 11	Tm 12	Man 13	Accomp 14
S_{nex}	(32)				S			Pr			C	(±Ad)	(±Ad)		
	(33)	PERMUTATION OF (32)													
	(34)				S^x/C^y			Pr				(±Ad)	(±Ad)		
S_{st}	(35)							S/Pr			C	(±Ad)	(±Ad)		
	(36)						Pr	S	(±O)		O	(±Ad)	(±Ad)		
	(37)						Pr	S	O			(±Ad)	(±Ad)		
S_{act}	(38)						S	Pr			C	(±Ad)	(±Ad)		
	(39)						S/Pr					(±Ad)	O		
	(40)	VARIANT OF (39)					S/Pr					(±Ad)	O		
	(41)						S	Pr			C	(±Ad)	(±Ad)		
	(42)						S/Pr/O					(±Ad)	(±Ad)		
S_{mo}	(43)					S		Pr			C	(±Ad)	(±Ad)		
	(44)			C^y		S^x		Pr				(±Ad)	(±Ad)		
	(45)					S	Pr		O			(±Ad)	(±Ad)		

Diagram IV (concluded)

CHAPTER IV

ENGLISH SUBJECT TAGMEMES: CLAUSE LEVEL EXPONENCE RULES

In this final chapter we will present the exponents or fillers of the English subject tagmemes, distinguishing filler types and describing their internal structure. Before listing the exponents for each tagmeme, we will comment generally on the contrastive features of grammatical levels, making a distinction between exponence rules which are recursive and exponence rules which are progressive; recursive rules fill clause level tagmemes with clause level syntagmemes, while progressive rules fill clause level tagmemes with phrase level syntagmemes. At the end of the chapter we shall comment briefly on word level rules, attempting to project this fragmentary grammar through the grammatical hierarchy.

4.1 The Distinction between Clause Level and Phrase Level

Exponence rules for the English subject tagmemes supply labels and then syntagmemic formulae for filler types. Leaving aside pronoun exponents, which are constrained by contexts larger than a single clause and therefore beyond our focus here, there are three kinds of syntagmemes which are exponents of English subject tagmemes: 1) pure phrase syntagmemes, 2) reduced clause syntagmemes, and 3) combinations of phrase and reduced clause syntagmemes. The basic distinction between clause level and phrase level syntagmemes is that the grammatical form (though not necessarily the grammatical meaning) of the tagmemes which make them up are different. At clause level the grammatical forms (or relations) include: subject, predicate, object, complement, and adjunct. At phrase level there is another set of grammatical forms: determiner, head, subjunct, restrictive, and appositive. Thus, at clause level we might have the string,

$$S_{ag}\ Pr_{act}\ Ad_{tm}$$

describing, for instance, the clause,

Dogs bark in the evening.

At phrase level, on the other hand, we might have the string

$$Sb_{ag}\ Sb_{qual}\ R_{rel}\ H_{gl}\ A_{rel}$$

describing, for instance, the phrase,

My poor barking dog, Fido,...

The phrase level string can be read, "subjunct as agent (my), subjunct as quality (poor), restrictive as relative clause-reduced (barking who is barking), head as goal (dog), appositive as relative clause reduced (Fido whose name is Fido). Thus, the major distinction between clauses and phrases is that they describe different sets of grammatical relations.

It must be emphasized, however, that the distinction between clauses and phrases is not that phrases are endocentric and clauses are exocentric constructions. As has been pointed out (3.2.3) clauses are often "headed" constructions, though we have chosen to call such clauses "subject-focused," "predicate-focused," "object-focused," etc. As we said earlier, clauses can be seen as subject-focused unless the subject has been reduced. However, once the subject has been reduced, the syntagmeme still remains a clause. For instance, given the clause,

John built the boat.

there is an agnate reduced clause with predicate focus (and hence reduced subject focus),

John's building the boat

or,

For John to build the boat

The phrasal paraphrase of this clause has very different grammatical form:

The building of the boat

or

John's building of the boat

or, perhaps,

The John-built boat.

In both the clause syntagmemes and the phrase syntagmemes above, the grammatical meanings *agent*, *act*, *goal* are present, but they have different grammatical forms in clauses and phrases. Thus,

Clause: For John to build the boat = Nexus:
$S_{ag}\ Pr_{act}\ O_{gl}$ (see 4.4.1.3)

Phrase: The building of the boat (by John) =
$D_{sp}\ H_{act}\ Sb_{gl}\ (Sb_{ag})$ (see 4.4.2)

A secondary distinction between clause level and phrase level syntagmemes is that some categories of grammatical meaning seem to appear only at one level. Though we can find agent, goal, act, location, time, and most other clause level categories at phrase level, there seems no phrase category for state: a distinctive feature of clauses seems to be the grammatical meaning of "to be" which we have called state.[1] On the other hand, some of the grammatical meanings of the English determiner have no apparent clause level manifestations. The phrase level category, quantity, as in

The two boys...

has no clause paraphrase. That is, we cannot say,

*The boys are two.

The rare,

We are six.

(said perhaps to a head waiter) might be an exception, requiring the category quant at clause level; however we would consider the sentence above to have undergone a context conditioned deletion of a phrase head, e.g. "people," in which case quant remains a phrase level constituent.[2]

Though levels may have some unique categories, we have not used this as a criterion for distinguishing levels. Contrastive grammatical form seems sufficient.

4.2 Distinctive Types of Subject Exponents.

As we mentioned above, there are three kinds of syntagmemes (excluding pronouns) which can be exponents of English subject tagmemes: 1) pure phrase syntagmemes, 2) reduced clause syntagmemes, and 3) combinations of phrase and reduced clause syntagmemes. Examples of these are,

(a) Two old friends (Pure phrase)

(b) Whoever are sleeping (Reduced clause)

(c) Two old friends who are sleeping (combination)

The two phrase constituents before the head in example (a) cannot, except apparently in an ad hoc way, be derived transformationally from clauses. They seem to manifest distinctive, phrase level tagmemes. That is, there are no clauses which might be considered paraphrases of (a):

*Friends are two.

*Friends are old.

However, the second tagmeme in,

(d) The barking dog

does have a clausal source:

The dog is barking.

The clause may be embedded whole in the phrase, as in

The dog who is barking.

or in reduced form,

The dog barking...

which may be permuted to

(d) The barking dog...

Hence we shall consider (d) an example of the combination of phrase and reduced clause syntagmemes which has undergone deletion and permutation. The rules for these processes will be discussed later (4.4.1.2).

4.3 Exponents of English Subject Tagmemes

The following is a list of exponents for each English subject tagmeme. After the symbol for each tagmeme, names for different phrase and reduced clause types are given; these are keyed to the readings rules for phrase and reduced clause syntagmemes, to be given later (4.5). Examples of each exponent type are given.

4.3.1 S_{ag}

Exponents: Phr$_1$ (agent head), Phr$_2$ (goal head), Phr$_3$ (quantity head), Phr$_4$ (quality head Relative Clause$_1$ (wh- agent clause) Relative Clause$_2$ (wh- goal clause)

Examples:

Phr$_1$: The old man (left).

Phr$_2$: My old dog (left).

Phr$_3$: One of them (left).

Phr$_4$: The truly brave (left).

R Cl$_1$: Whoever was in the room (left).

R Cl$_2$: Whoever was called (left).

4.3.2 S_{prop}

Exponents: Phr_5 (proposition head)

Proposition Clause (that - clause)

Examples:

Phr_5: The idea that John is frightened (amuses me).

Prop Cl: That John is frightened (amuses me).

4.3.3 S_{in}

Exponents: Phr_2, Phr_3, Phr_5, Phr_6 (instrument head)

Relative $Clause_3$ (wh- instrument clause)

Examples:

Phr_2: My key (opened the door).

Phr_3: Two of the keys (opened the door).

Phr_5: The idea that John is frightened (impresses me).

Phr_6: The key with which I opened the door (opened the gate).

R Cl_3: Whatever I opened the door with (opened the gate).

4.3.4 S_{gl}

Exponents: Phr_1, Phr_2, Phr_3, Phr_4, Phr_6

Relative $Clause_1$, Relative $Clause_2$, Relative $Clause_3$

Examples:

Phr_1: The old man (was found).

Phr_2: My old dog (was found).

Phr_3: One of them (was found).

Phr_4: The truly brave (were found).

Phr_6: The key with which I opened the door (was found).

R Cl_1: Whoever was in the room (was found).

R Cl_2: Whoever was called (was found).

R Cl_3: Whatever I opened the door with (was found).

4.3.5 S_{tm}

Exponents: Phr_7 (time head), Relative Phrase_1 (tm relator-axis propositional phrase)

Relative Clause_4 (wh- time clause)

Examples:

Phr_7: The morning (is the time to go).

R Phr_1: From June to September (is the time to go).

R Cl_4: Whenever John is ready (is the time to go).

4.3.6 S_{loc}

Exponents: Phr_8 (location head), Relative Phrase_2 (loc relator-axis prepositional phrase)

Relative Clause_5 (wh- location head)

Examples:

Phr_8: Chicago (is too far).

R Phr_2: From here to Chicago (is too far).

R Cl_5: Wherever John is staying (is too far).

4.3.7 S_{qual}

Exponent: Phr_9 (quality head)

Example:

Phr_9: His eagerness to please (amuses me).

4.3.8 S_{nex}

Exponent: Nexus Clause

Example:

Nex Cl: For John to be frightened (is ridiculous).

4.3.9 S_{st}

Exponent: Predicate $Clause_1$ (state head)

Example:

Pred Cl_1: Being cold (is no fun).

4.3.10 S_{act}

Exponents: Phr_{10} (action head)

Predicate $Clause_2$ (action head)

Examples:

Phr_{10}: John's love of sports (is obsessive).

Pred Cl_2: John's loving sports (pleases his mother).

4.3.11 S_{mo}

Exponents: Phr_{11} (motion head)

Predicate $Clause_3$ (motion head)

Examples:

Phr_{11}: The dance with Mary (was a privilege).

Pred Cl_3: Dancing with Mary (was a privilege).

4.4 The Internal Structure of Exponents

As we saw in Chapter II, after clause level tagmemes have been read via clause level reading rules (and optionally permuted or conjoined, operations we have not described in detail here), they are given symbols for particular exponents; for example, S_{ag} can be given the exponent Phr_1, which can be represented as S_{ag}: Phr_1. A full representation of the tagmeme (excluding aspect D, which has not been properly formulated) would include the grammatical form and meaning and all the potential exponents: S_{ag}: $Phr_1/Phr_2/Phr_3/Phr_4/$ R Cl_1/ R Cl_2. Stated as rewrite rules, this would take the form:

$$S_{ag} \rightarrow \begin{bmatrix} Phr_1 \\ Phr_2 \\ Phr_3 \\ Phr_4 \\ R\ Cl_1 \\ R\ Cl_2 \end{bmatrix}$$

In the generative process, these labels for clause level exponents are given readings which describe their internal structure, either as phrase level syntagmemes or reduced clause level syntagmemes. These readings, like those for clause level syntagmemes, are readings across rows in matrices in which the rows are marked by the exponent label (e.g. Phr_1) and the columns by the categories of grammatical meaning; grammatical forms, carrying signs to mark optional tagmemes, fill the cells. Such a matrix can be found in section 4.5 below.

4.4.1 Reduced Clause Allo-syntagmemes as Exponents of English Subject Tagmemes

The exponence rules include four kinds of reduced clauses. (See 3.2.3 for discussion of reduced clause) These are relative clauses (R Cl), proposition clauses (Prop Cl), nexus clauses (Nex Cl), and predicate clauses (Pred Cl). These reduced clauses are derived from the clause level matrix (diagram III) by rather simple operations. That is, we can define an operation on an entire matrix which produces a derived matrix.[3] If we represent the matrix in Diagram IV as M_1, the operation "reduction to Prop Cl" as P, and the derived matrix as M_1P, then we can describe reading the exponent Prop Cl as any reading across rows in M_1P. Likewise, reduction to relative clauses, nexus clauses, and predicate clause can be presented as M_1R, M_1N, and M_1Pr, respectively. Let us now briefly define these operations.

4.4.1.1 Proposition Clauses as Exponents of English Subject Tagmemes

This operation is simply, "add THAT as initial symbol in each row." This operation yields, for example, the reduced proposition clause

$$\text{THAT } S_{ag} Pr_{mo} \pm O_{gl} \pm Ad_{loc} \pm Ad_{tm} \pm Ad_{man} \pm Ad_{accomp}$$

(That John walked five miles in the woods today slowly with Mary) as a reading of the first row of M_1P. All of the rows in M_1 can be similarly read. However, some permutations of readings from M_1 cannot undergo this operation and yield correct English clauses. If $\underline{+}Ad_{man}$ had been permuted to clause initial position in the example above, the operation for Prop Cl would yield

*That slowly John walked five miles in the woods....

Hence, clause reduction restricts permutation rules, although the tagmemes in reduced clauses can be permuted. For example, $\underline{+}Ad_{man}$ can be permuted to the position preceding Pr_{mo} in the Prop clause above:

That John slowly walked five miles....

Clearly, permutation rules must follow clause reduction.

4.4.1.2 Relative Clauses as Exponents of English Subject Tagmemes

The operation that yields the derived matrix M_1R is the most complex clause reduction operation, for it requires the addition of a relative marker (wh-) to a tagmeme, the obligatory permutation of this tagmeme to clause initial position, and sometimes the change of the sign (+ $\underline{+}$) before the permuted tagmeme and before the tagmeme Pr_{st} (if this tagmeme is read). The relative marker (wh-) is a constraint on the exponent of the tagmeme. Thus adding wh- to S_{ag} might be marked S_{ag}:wh-, which, in effect, tells the grammar that at word level the exponent of this tagmeme must be who or whoever. If wh- is added to Ad_{ag}, the word level exponent will be by whom. If wh- is added to Ad_{in}, the word level exponent will be with which, etc.

The change of sign before the tagmeme marked by wh- and Pr_{st} can be described only in a more complete grammar than this. We can only sketch out the process. In phrases like,

The man who was singing
The dogs who were barking
The boy who is my brother

the relative word and the copula are clearly optional, for we can have the following phrases, which are variations of the three above:

The man singing
The dogs barking
The boy, my brother,

Furthermore, the tagmemes which are restrictive (not appositive) can be permuted to

The singing man
The barking dogs

In order to generate strings of the form Who+BE+V-ing at clause level we will, following Harris[4], probably read these strings from a base structure

$S\ Pr_{st}\ C_{act/mo}$
e.g. He (is) (singing)

where the past and present participle preceded by BE are seen as exponents of the complement, not the predicate. Thus the sign before the subject (marked by wh-) and the predicate (state) can be changed from + to ±. This is apparently not a completely ad hoc analysis, for the conjoining rules seem to require this identification of

Pr (Be) and C(V-ing) as separate clause level tagmemes:

He is singing and dancing.

However, this is to forecast, hastily, decisions about matters beyond the scope of this partial grammar.

Hence we must assume an operation on M_1 which yields M_1R as a source of embedded relative clauses in the noun phrase. This is a vital but reasonable assumption.[5]

4.4.1.3 Nexus Clauses as Exponents of English Subject Tagmemes

The operation which yields M_1N from M_1 is, fortunately, not difficult to describe. It operates identically on all the syntagmemes represented by M_1, and, like the operation for Prop Cl but unlike that for R Cl, it seems to affect only the surface structure of the clause syntagmemes. The rule is simply, "add the marker +For to the subject, the marker +To to the predicate, and change all signs to ± except that on the predicate, which remains +." The only qualification is that if the clause has a dummy subject (there), the marker +For is added to the dummy subject:

There is no doubt that John is frightened.

For there to be no doubt that John is frightened.

4.4.1.4 Predicate Clauses as Exponents of English Subject Tagmemes

We said earlier (3.2.3) that non-reduced English clauses seem to be subject-focused clauses. Predicate clauses are clearly predicate-focused:

His singing the song over and over bothered me.

Like nexus clauses, all the tagmemes except the predicate tagmeme are optional:

Singing bothered me.

The rule for deriving M_1Pr from M_1 is "add marker -'s to subject and marker -ing to predicate (these markers, like wh-, +For, and +To are word level exponence constraints and not actual phonemic representations of morphemes.), and change all signs to ± except that on the predicate, which remains +."

There is a major qualification on this rule: not all the syntagmemes in M_1 can undergo this operation. For instance, we must not generate strings like,

*The idea that John is frightened's amusing me.

*Kindness's being killed by the cat.

*Walking's being fun.

Like the operation for deriving relative clauses, this operation is constrained by the deep structure of syntagmemes. However, unlike the rules for deriving relative clauses, these for deriving predicate clauses are statable within the limits of this partial grammar. The reason for this is that the constraints from the deep structure here seem to involve only the subject. Only clauses with concrete subjects (i.e. S_{ag}, S_{in}, and S_{gl}) seem to allow this derivation:

John's running in the park every morning...

The stone's striking Bill...

Bill's being hit by John...

Some people, when asked, seem to accept some instances of S_{tm} and S_{loc}:

Sunday's passing quickly...

Ann Arbor's being a nice place to live...

Hence, we shall limit the operation Pr on M_1 to rows with S_{ag}, S_{in}, and S_{gl} (and marginally S_{tm} and S_{loc}).

Just as the derivation of a Prop Clause from M_1 restricted permutation rules (4.3.2), the derivation of a Pred Clause restricts exponence rules. Subjects of Pred Clauses can have as exponents only pure phrases (not reduced clauses or phrases containing restrictive or appositive clauses). That is, the grammar must not generate clauses like either of the following (both S_{ag}):

*Whoever was in the room's killing Bill...

*The man who sold cigarettes's leaving town...

But it will allow pure phrases like,

The King of England's signing the paper...

Both of these constraints (i.e. limiting the operation to certain rows of M_1 and restricting exponence rules) can be included without difficulty within the model proposed here, however. That is, they complicate the grammar but do not seriously challenge it.

4.4.1.5 Summary of Reduced Clause Allo-syntagmemes

General formulae for reduced clause allo-syntagmemes are listed below. These formulae represent only gross abbreviations of rows in reduced clause matrices. Ellipses in the formulae, preceded by a sign (+ or $\pm$), represent all the remaining tagmemes, obligatory and optional, in that syntagmeme type.

Prop Cl = THAT + Cl

Rel $\text{Cl}_1 = \text{S}_{\text{ag}}$:wh- +....

Rel $\text{Cl}_2 = \text{S}_{\text{gl}}$:wh- +....

Rel $\text{Cl}_3 = \text{S}_{\text{in}}$:wh- +....

Rel $\text{Cl}_4 = \text{S}_{\text{tm}}$:wh- +....

Rel $\text{Cl}_5 = \text{S}_{\text{loc}}$:wh- +....

(for all Rel Cl if Pr_{st}, then $\pm\text{S}_{\text{x}}$:wh- $\pm\text{Pr}_{\text{st}}$ +C...)

Nex Cl = $\pm$ FOR S + TO Pr $\pm$....

Pred $\text{Cl}_1 = \pm\ \text{S}_{\text{ag/in/gl}}:\underline{\text{-'s}}\ +\text{Pr}_{\text{st}}$:Ving $\pm$....=

Pred $\text{Cl}_2 = \pm\text{S}_{\text{ag/in/gl}}$:-'s $+\text{Pr}_{\text{act}}$:Ving $\pm$....

Pred $\text{Cl}_3 = \pm\text{S}_{\text{ag/gl}}$:-'s $+\text{Pr}_{\text{mo}}$:Ving $\pm$....

4.4.2 Pure Phrase Syntagmemes as Exponents of English Subject Tagmemes

As we stated above, phrase syntagmemes are distinguished from clause syntagmemes because the constituents of each are related to each other in different ways. That is, each have different grammatical forms: clauses include the relations subject, object, predicate, complement, and adjunct; phrases include the relations determiner, subjunct, head, restrictive, appositive. Only the first three of these occur in pure phrase syntagmemes. The relations restrictive and appositive are always manifested by reduced clause syntagmemes, and hence syntagmemes including these forms are combinations of phrase and reduced clause syntagmemes. In this section we will describe only the pure phrase syntagmemes.

The forms determiner (D), subjunct (Sb), and head (H) are, like other aspects of grammatical units, defined operationally by the grammar. We can, however, describe a number of the distinctive features of each.

4.4.2.1 *Determiners*

Determiner tagmemes are obligatory in the phrase syntagmeme, unlike subjunct tagmemes, which are optional. Furthermore, they are not permutable or conjoinable, while subjunct tagmemes, as we shall see, are. There are three categories of grammatical meaning with which the form determiner may occur; these are all apparently distinctive phrase level categories, though they are not limited to the subject phrase. These categories are specification (*spec*), emphasis (*emp*), and quantity (*quant*). The function of the category *spec* is primarily anaphoric; it marks the relation of a phrase to another phrase or clause (either embedded or outside the phrase) or to the situation outside the utterance. This category *spec* may occur as the determiner (D_{spec}) as in,

The book on the table....

Here D_{spec} ('the') relates the phrase, 'the book' to the subject (lost in the surface structure) of the embedded clause (R Cl_2)'which is on the table.' Likewise, in the phrase,

The last man....

where D_{spec} ('the') relates the phrase 'the man' to the subject of an embedded, permuted clause (R Cl_1) 'who was last.' Sometimes the anaphoric reference is beyond the domain of the sentence, as in

The man removed his hat.

Here we assume a context in which man is identified. The word level exponents (See 4.6) of D_{spec} are the, a, and Ø.

The category spec may occur as subjunct or as head, as well as determiner, as in the following sentences:

Both these books were stolen. (Sb_{spec})

This is not clear. (H_{spec})

If the obligatory determiner is not D_{spec} it may manifest either the category emp (emphatic) or the category quant (quantity). The category emp has word level exponents like each, every, just, any, only. Examples of D_{emp} are

Just two books....

Any two books....

Only my book....

The category emp may also occur as subjunct or as head, as in the following phrases:

The only book.... (Sb_{emp})

Any of the books.... (H_{emp})

Finally, determiner may occur as quant (D_{quant}) as in the following phrases:

All sixteen books.... (D_{quant} and Sb_{quant})

Some old books....

Several books....

The category quant may also occur as Sb_{quant} and H_{quant} as in

Only two books.... (Sb_{quant})

Only two of them.... (H_{quant})

It may be noted that exponents of D, Sb, and H at the word level may be identical in surface form. For instance, a word manifesting quant may be D, Sb, or H:

Many small books.... (D_{quant})
The many books.... (Sb_{quant})
Many of them.... (H_{quant})

This is no different, however, from the identical word level manifestations of many clause level tagmemes:

The book burned. (S_{gl})
I burned the book. (O_{gl})
Love is wonderful. (S_{act})
We love springtime. (Pr_{act})

This merely illustrates what we argued in Chapter I: grammatical forms (aspect A) are not predictable from lexical forms (aspect C) or vice versa.

4.4.2.2 Subjuncts

Subjuncts, unlike determiners, are optional, permutable, and conjoinable. As we have seen in 4.4.2.1 above, subjuncts may occur as Sb_{spec}, Sb_{emp}, and Sb_{quant} as well as Sb_{ag}, Sb_{gl}, Sb_{qual}, and others. Examples of the latter are:

Both my books.... (Sb_{ag})
The flower's smell.... (Sb_{gl})
A major decision.... (Sb_{qual})

The tagmemes Sb_{emp} and Sb_{quant} may be permuted to phrase initial (pre-determiner) position:

Only the book.... ($Sb_{emp}D_{spec}H_{gl}$)

Both the books.... ($Sb_{quant}D_{spec}H_{gl}$)

Other tagmemes, Sb_{ag} and Sb_{gl}, may be permuted to a position following the head:

Both books of mine.... ($D_{quant}H_{gl}Sb_{ag}$)

The smell of the flower.... ($D_{spec}H_{act}Sb_{gl}$)

Because permutation rules precede exponence rules, we will not generate such strings as

*Each the books....

*Some the books....

or

*Both books my....

*A decision of major....

That is, the exponents of a permuted tagmeme are not necessarily the same as those of the same tagmeme when it has not been permuted. While all and both are word level exponents of the permuted Sb_{quant}, some and many are word level exponents of Sb_{quant} when it has not been permuted. We will discuss the derivation of the embedded phrase 'of mine' in section 4.4.4 below

Likewise, subjunct tagmemes may be repeated (conjoined). Like clause level conjoining, conjoining at phrase level is not a surface operation but requires K-equivalence, i.e. equivalence of both grammatical form (surface structure) and grammatical meaning (deep structure). For instance, conjoining rules will allow these strings:

Both John's and Bill's books.... ($D_{quant}Sb_{ag}$ and $Sb_{ag}H_{gl}$)

Some but not all the students.... (Sb_{quant} but not $Sb_{quant}D_{spec}H_{ag}$)

The conjoining rules will not permit these strings:

*My and the flower's smell.... (Sb_{ag} and $Sb_{gl}H_{act}$)

*Two and my books.... (Sb_{quant} and $Sb_{ag}H_{gl}$)

One category which has been considered by others (e.g. Liem, 1966, p. 84) to be part of the pure phrase (order or rank) is here considered a restrictive, derived constituent of the phrase. That is, in the phrase,

The last two boys....

the word, 'last,' is considered to be a manifestation of the tagmeme R_{rel} permuted from a reduced clause,

The two boys who were last....

In the reduced clause 'who were last,' 'last' is a manifestation of the clause-level tagmeme C_{qual}. Evidence for this interpretation is that order words operate as restrictives, which affect the selection of D_{spec}. The following examples should make this clear:

The major decision.... (non-restrictive)

A major decision....

The last man.... (restrictive)

*A last man....

This distinction will be discussed further in section 4.4.3 below.

4.4.2.3 Heads

The head tagmeme in the pure phrase is obligatory, conjoinable, but not permutable. In a sense, it is the axis around which phrase level permutations move. The head of a phrase can, as we have seen, occur as *spec*, *emp*, *quant*, *ag*, *gl*, *act*, *in*, *loc*, *tm*, *prop*, *qual*, *mo*, and probably most of the other categories (except *st*) as well. The following are illustrations of various kinds of phrase heads:

H_{spec}: All *that* (is finished).

H_{emp}: *Each* of them (must go).

H_{quant}: Only *sixteen* of them (can come).

H_{ag}: The old *man* (looked at me).

H_{gl}: My *book* (is on the table).

H_{act}: His *fear* of death (is degrading).

H_{in}: The *key* (with which I opened the door) (got lost).

H_{loc}: The *center* of Ann Arbor (is rather shabby).

H_{tm}: This past *week* (has been hectic).

H_{prop}: Both his *ideas* (are wrong).

H_{qual}: His *anxiety* for news (leads him to buy four newspapers).

H_{mo}: His *walking* of the dog (takes nearly an hour).

We have seen (4.4.2.2) that both agent and goal may be subjuncts of a head which is H_{act}, H_{qual}, and H_{mo}, and that these subjuncts have different forms preceding and following the head. Let us examine this more carefully. $H_{act/mo}$ or H_{qual} will have as word level exponents nominalized forms of verbs and adjectives, respectively, or else -ing forms of verbs and non-nominalized adjectives. The differences in these exponents will in some cases be distinguished only at word level.

(a) His construction of the building.... ($Sb_{ag}H_{act}Sb_{gl}$)

(b) His constructing of the building.... ($Sb_{ag}H_{act}Sb_{gl}$)

(c) His eagerness to serve.... ($Sb_{ag}H_{qual}R_{nex}$)

(d) The eager.... ($D_{spec}H_{qual}$)

Both (a) and (b) have identical phrase level structure and both are exponents of S_{act}. Examples (c) and (d), on the other hand, have somewhat different phrase level structure, for (c) is an exponent of S_{qual} while (d) is an exponent of S_{ag} or S_{gl}. Hence (c) and (d) will be distinguished in the phrase level exponence rules, while (a) and (b) will not.

Notice that the subjuncts of (a) and (b) are not derived from embedded, reduced relative clauses. That is, there is no phrase

*His construction which was of the building....

from which to derive (a) or (b). Hence (a) and (b) are pure phrase syntagmemes, not combined phrase and clause syntagmemes (as is, for instance, (c) which includes a restrictive nexus clause).

The subjuncts in pure phrase syntagmemes like (a) and (b) occur both preceding and following the head. It is possible that both Sb_{ag} and Sb_{gl} may occur preceding $H_{act/mo}$, though there are some constraints here which await a more precise description of clause level predicate categories than has been given here. When both Sb_{ag} and Sb_{gl} precede H_{act}, then Sb_{ag} precedes Sb_{gl}, which we would specify in phrase level permutation rules. An example of this order is

John's tree pruning (leaves something to be desired).

However, if Sb_{gl} is to be specified (and not generic), the tagmeme must follow the head as a relative phrase (see 4.3.4) marked by _of_:

John's pruning of that tree (leaves something to be desired).

When both Sb_{ag} and Sb_{gl} follow the head, there is no fixed order but each subject is a relative phrase marked by _by_ and _of_ respectively:

The pruning of that tree by John....

The pruning by John of that tree....

The distinction of Sb_{ag} and Sb_{gl} is always clearly marked when both occur, but the choice of relative phrase marker (preposition) depends upon the predicate category chosen. Compare the phrases above with the following:

John's love of Mary....

The love of John for Mary....

Here the tagmeme Sb_{gl} is marked by _of_ if Sb_{ag} precedes the head and by _for_ if both follow the head. If only one of these subjunct tagmemes occurs following the head, then the phrase is often ambiguous:

The love of John....

Here 'of John' may manifest either Sb_{ag} or Sb_{gl}.

The categories _tm_ and _loc_ may also occur as subjuncts with $H_{act/mo}$. The following phrases illustrate these other subjuncts in action headed subject phrases. None of them are derived from embedded relative clauses in these instances.[7]

The noon conference.... ($D_{spec}Sb_{tm}H_{act}$)

The conference at noon.... ($D_{spec}H_{act}Sb_{tm}$)

The Washington conference.... ($D_{spec}Sb_{loc}H_{act}$)

The conference in Washington.... ($D_{spec}H_{act}Sb_{loc}$)

Many abstract nouns are exponents of $H_{act/mo}$ (e.g. love, life, death, speech, residence, contest, etc.) or H_{prop} (e.g. fact, idea, notion, etc.) Some abstract nouns may be exponents of either (e.g. hope, fear, statement, claim, etc.) Phrases with H_{act} and H_{prop} are clearly contrastive in structure, however.

The death of a poet.... ($D_{spec}H_{act}Sb_{ag}$)

The fact that John is frightened.... ($D_{spec}H_{prop}A_{prop}$)

John's fear of Bill.... ($Sb_{ag}H_{act}Sb_{gl}$)

John's fear that he will be killed.... ($Sb_{ag}H_{prop}A_{prop}$)

Other abstract nouns are nominalizations of adjectives which are exponents of H_{qual} (e.g. anxiety, eagerness, sincerity, etc.) This tagmeme is frequently followed by A_{nex}, as in the following examples:

His anxiety to leave....

His anxiety for Bill to leave....

His eagerness for news....

His eagerness for news to break....

In all these examples, the phrase structure is

$Sb_{ag}H_{qual}A_{nexus}$

and all are exponents of the clause level tagmeme S_{qual}.

4.4.3 Combined Phrase and Clause Syntagmemes (Recursive Phrase Structure)

So far we have discussed two kinds of exponents of English subject tagmemes, reduced clause syntagmemes and pure phrase syntagmemes. A third general kind of subject exponent is the combined pure phrase and reduced clause syntagmeme. As we saw above, pure phrase tagmemes include only the grammatical relations determiner, subjunct, and head. The two other grammatical relations in phrase level syntagmemes are restrictive (R) and appositive (A) which are always manifested by reduced clauses.

Restrictive and Appositive tagmemes can probably be distinguished only by lexical and phonological criteria. An appositive tagmeme follows the head and is separated from the head by comma intonation.[8] It may not be permuted to a pre-head position, as some restrictive clauses may. However, there seems no purely grammatical way to dinstinguish the following two clauses, the first containing a restrictive reduced clause and the second an appositive reduced clause:

(a) The rocket which requires two years to build is worthless.

(b) The rocket, which requires two years to build, is worthless.

Furthermore, the determiner does not help to distinguish these sentences, as the following sentences show:

A rocket which requires two years to build is worthless.

A rocket, which requires two years to build, is worthless.

> Any rocket which requires two years to build is worthless.
>
> Any rocket, which requires two years to build, is worthless.

There are obvious differences in the meaning of these sentences and obvious differences in their phonology. There are at least two approaches to the possible distinguishing of them grammatically. The first is somewhat circular: we might posit an optional appositive marker tagmeme which would follow the subject in appositive clauses; this tagmeme would be manifested by words and phrases like 'incidentally,' 'by the way,' etc. This solution seems ad hoc, however, for it does not occur in all appositive clauses, e.g.

> *The idea that John, incidentally, is frightened....

The second possible approach offers better hope. In example (a) above, the determiner 'the' is either anaphoric or generic. If it is anaphoric, its reference is to the subject of the embedded clause, 'which.' In example (b), 'the' is also either anaphoric or generic, but if it is anaphoric, its reference is to a context (linguistic or non-linguistic) outside of the sentence. If we can devise a descriptive device to specify the domains of anaphoric references, then we might formally explain the difference between (a) and (b).[9]

Even though we cannot fully explain the difference between restrictive and appositive embedded clauses, it seems correct to generate two different tagmemes as phrase constituents, R and A. As we shall see, some, but not all, R tagmemes can be permuted to a pre-head position. An alternate, perhaps more consistently tagmemic approach would be to describe all phrase constituents short of full clauses as subjuncts manifesting categories like tm, act, loc, man, etc. Then the phrase,

The man on the couch....

would be described as,

$$D_{spec} H_{ag} Sb_{loc}$$

and not as a reduction of

The man who is on the couch....

$$D_{spec} H_{ag} R_{RCl}$$

However, then the phrase,

The man sleeping peacefully on the couch....

would be described

$$D_{spec} H_{ag} Sb_{act} Sb_{man} Sb_{loc}$$

This description seems wrong, for 'peacefully' and 'on the couch' are clearly adjuncts of the predicate 'sleeping' and its deleted subject rather than subjuncts of the head 'man.'

The description of phrase structure requires both derived (embedded) and non-derived (base) constituents, and the issue here is not between one or the other; it is rather a decision about which tagmemes are which. The choice here is to derive a string like

The sleeping man....

from the combined phrase and clause syntagmeme,

The man who is sleeping....

following, in this instance, the lead of Chomsky and Lees.

Hence, we shall posit the phrase level tagmemes R and A which are always filled by either a reduced prop clause (PropCl), a reduced nexus clause (NexCl) or a reduced relative clause (RCl). Reduced predicate clauses are never exponents of R or A (phrase level tagmemes) but only of S_{st}, S_{ag}, or S_{mo} (clause level tagmemes).

As a phrase constituent, PropCl occurs only as A_{PropCl}, following H_{prop} to yield phrases like,

The idea that John is frightened....

His fear that war may come....

Any statement that it is inevitable....

A phrase containing H_{prop} may also include A_{NexCl} as in

His desire for John to join the Peace Corps....

His wish to avoid war....

My hope for him to succeed....

However, a phrase with H_{prop} may not include both A_{PropCl} and A_{NexCl}, even as conjoined tagmemes:

*My hopes that John will graduate and for peace to come....

A_{PropCl} and A_{NexCl} are repeatable, though:

My hopes that John will graduate and that peace will come....

My hopes for John to graduate and for peace to come....

As we shall see below, R_{rel} and A_{rel} may occur with any head tagmeme, including H_{prop} when it is followed by A_{PropCl} or A_{NexCl}, as in the following:

The hope which I have often stated that peace will come....

The hope for peace to come, which I have often stated....

The phrase level exponent NexCl occurs as A_{NexCl} with H_{prop}, as we have just seen, and with H_{qual}, as in

His eagerness for you to leave....

His anxiety for news to come....

It occurs as R_{NexCl} with other head tagmemes. Here are some examples:

The man for you to watch.... (H_{ag})

The key for you to open this door with.... (H_{in})

The idea for you to remember.... (H_{prop})

The life for you to lead.... (H_{act})

Relative clauses occur as R_{rel} and A_{rel} with all head tagmemes, e.g.:

The man who has won.... (H_{ag})

My dog, which was run over,.... (H_{gl})

The idea you approve of.... (H_{prop})

The love of life which John displays.... (H_{act})

The dancing which annoys you.... (H_{mo})

Furthermore, as we have seen, when the tagmeme R_{rel} has the form

$S\ Pr_{st} C_{qual/act/mo}$

(e.g. who is sleeping, which is cold, that is singing, etc.)

and the optional subject and predicate are not read (or are deleted), then the complement C can be permuted to a position before the head following any subjunct or determiner, as in,

The sleeping man....

The cold key....

The singing canary....

The permutation rules must include the instructions that C_{qual} is obligatorily permuted, to avoid,

*The key cold....

and that, if the embedding is repeated, qual precedes act or mo, to avoid,

*The singing cold canary....

A much more detailed analysis of the clause is necessary before the numerous constraints on permutations of reduced restrictive clauses can be described in more detail. For instance, the distinction that Fillmore (1967, p. 7-8) discusses between affiziertes Objekt, e.g.

John ruined the table.

and effiziertes Objekt, e.g.

John built the table.

would probably be seen here as a difference in the predicate category rather than the object category. Nevertheless, this distinction does affect the permutability of C_{act} in a restrictive clause. We can have both of the following, where the head is the affectum:

A skillfully stolen car....

A stolen car....

However, when the head is the effectum, only one of these phrases may occur:

A skillfully built car....

*A built car....

Many more such distinctions will be necessary to explain restrictive embedding in any thoroughgoing way.

4.4.4 Relative (Prepositional) Phrases

The exponents of some clause level subject tagmemes may be relative phrases, e.g.

In the well is where I found it.

S_{loc}:$RPhr_2$ (loc relator-axis phrase)

At noon will be too late.

S_{tm}:$RPhr_1$ (Tm relator-axis phrase)

Relative phrases may also, of course, occur as exponents of many other clause level and phrase level tagmemes. Relative phrases, like relative clauses, are marked by a form which identifies an underlying grammatical category. That is, relative words marking relative clauses replace and identify a particular category: who marks S_{ag} or S_{gl}, whom or which marks O_{gl}, with which marks Ad_{in}, where marks Ad_{loc}, etc. Relative words (prepositions) marking relative phrases do not replace part of the phrase but are added (like that to a prop clause) to any phrase syntagmeme and identify the underlying grammatical category manifested by the head of the phrase. This marking of categories is not unambiguous, however, for some markers (prepositions) occur with more than one category and some categories are marked by many relative words.[10] The following is a partial listing:

Category	Relative Phrase Marker
Agent	of, by (the love of John for Mary, the building of the school by John)
Goal	of, for, to (see examples above, a call to John)
Instrument	with, by (construction with a hammer, manufacturing by hand)
Location	at, in, to, toward, beside, under, etc. (at home, in the shade, beside me, etc.)
Time	at, in, before, after, during, etc. (at noon, before dinner, during the war, etc.)
Manner	with, under (with care, under pressure, with speed)
Proposition	of, about (of the idea that he introduced, about the plan for us to finish)
Action	of (message of love, matter of believing)
Quality	of (feeling of anxiety, shade of red)

As additional categories are established at both phrase and clause level, additional relative clause and phrase syntagmemes will be identified. For example, the category source was mentioned as a probable clause level category not yet established in this grammar (see 3.3 above). This category seems to be marked by of and from:

I built the house of bricks.

I bought the house from John.

At phrase level this category is probably the one that follows certain H_{ag} and H_{gl} when these heads refer to various collections or classifications of people or things:

A book of poems....

A group of children....

A pound of rice....

A brace of partridges....

The relative phrase syntagmemes can be derived from the matrix of phrase syntagmemes (4.5) via a matrix operation, just as relative clauses are derived from the matrix of clause syntagmemes. In this operation the relative phrase marker (preposition) will agree in category with the head of the phrase. When a category is marked by many prepositions, the selection will probably have to be explained under entries for particular morphemes in the lexicon, though, as we shall see below, context sensitive exponence rules at the word level will cut down the range of selection, just as, for instance, the exponence rules for Ad_{ag} will correctly select the preposition by for the relative agent phrase (Ad_{ag}: by+AgPhrase).

4.5 Phrase Level Reading Rules for Exponents of English Subject Tagmemes

When the exponent of a subject tagmeme is a reduced clause, the grammar will cycle back to a clause level reading from one of the reduced clause matrices discussed in 4.4.1 above. When the exponent is a pure phrase or combined·phrase and reduced clause syntagmeme, the grammar will progress to a phrase level reading from the matrix below. Like the clause level matrix, this one is merely illustrative, sufficient only to give readings for the types of phrases which we have discussed here. (Examples follow the matrix)

The phrase level matrix (Diagram V) follows the same general format as the clause level matrix (Diagram IV). The rows describe twelve phrase syntagmemes which are exponents of English subject tagmemes. The columns are headed by categories of grammatical meaning (aspect B) which are operative in these syntagmemes. In the cells a are symbols for grammatical forms or relationships (aspect A) found at phrase level. These are D 'determiner,' Sb 'subjunct,' H 'head,' R 'restrictive,' and A 'appositive.' Any reading of R or A requires the grammar to cycle back to the reduced clause matrices discussed in (4.4.1). Tagmemes (except D) in a reading, like those at clause level, may be repeated in projected phrase level conjoining rules (not included here). Also tagmemes may be permuted (some obligatorily) in projected phrase level permutation rules (also not included here but discussed below). Optional tagmemes selected for the illustrations will retain the ± symbol to show that they are options. Options not selected will be shown in parentheses. Hence an empty cell indicates that the particular category does not occur in that syntagmeme as a phrase level constituent, though it may occur as a constituent of an embedded relative clause or relative phrase.

Phrase Level Reading Rules for Exponents of English Subject Tagmemes[11]

Diagram V

Categories of Grammatical Meaning:

		Spec	Emp	Quant	Qual	Ag	In	Gl	Prop	Mo	Act	Loc	Tm	Man	Nex	Rel
Phr_1(ag head)	(1)	D	±Sb	±Sb	(±Sb)	H	±Sb								(±R)	±R(±A)
Phr_2(gl head)	(2)	D	(±Sb)	(±Sb)	(±Sb)	±Sb		H							(±R)	±R ±A
Phr_3(quant head)	(3)	±Sb	D	H		±Sb									(±R)	(±R±A)
Phr_4(qual head)	(4)	D	(±Sb)		H									±Sb		(±R±A)
Phr_5(prop Head^a)	(5)	D	(±Sb)	(±Sb)	(±Sb)	(±Sb)			H/±A							±R(±A)
Phr_5(prop head^b)	(6)	D	(±Sb)	(±Sb)	(±Sb)	±Sb			H						±A	(±R±A)
Phr_6(in head)	(7)	D	(±Sb)	(±Sb)	(±Sb)	±Sb	H	±Sb							(±R)	(±R±A)
Phr_7(tm head)	(8)	D	±Sb	(±Sb)	(±Sb)	(±Sb)							H		±R	±R(±A)
Phr_8(loc head)	(9)	D	(±Sb)	(±Sb)	(±Sb)	(±Sb)						H			±R	±R(±A)
Phr_9(qual head)	(10)	D	(±Sb)		H	±Sb									±A	(±R±A)
Phr_{10}(act head)	(11)	D	(±Sb)		(±Sb)	±Sb		±Sb			H			(±Sb)	(±R)	(±R±A)
Phr_{11}(mo head)	(12)	D	(±Sb)	±Sb	(±Sb)	(±Sb)		(±Sb)		H				(±Sb)	(±R)	±R(±A)

Examples of Syntagmemes in Diagram V

(1) The only two boys with knives in the room....

(2) My old dog, Fido,....

(3) Just these two of mine....

(4) The truly brave....

(5) The old idea that love conquers all....

(6) My wish for the war to end....

(7) My key to the door....

(8) The only day to go when I am free....

(9) A good place to visit....

(10) His eagerness to please....

(11) John's love of sports....

(12) The two dances with Mary....

4.6 Projection to Word Level

We can merely illustrate in this partial grammar how the generative process might be projected down through the grammatical hierarchy.[12] We will describe the possible derivation of the grammatical structure of a single tagmeme from clause level to word level. The clause is a rather simple one:

His two notebooks are lost.

We will assume that we enter the grammar after the sentence level has generated as an exponent $\text{Clause}_{\text{goal focus}}$. This then is given a reading at clause level (Diagram IV):

(1) $\text{Pr}_{\text{st}}\text{S}_{\text{gl}}\text{C}_{\text{qual}}$

In the permutation rules the subject tagmeme would obligatorily be permuted to initial position:

(2) $S_{gl}Pr_{st}C_{qual}$

As there are no clause level conjoinings here, we proceed to exponence rules. The exponents for S_{gl} are Phr_1, Phr_2, Phr_3, Phr_6, RCl_1, RCl_2, RCl_3. Phr_2 (gl head) is selected. The formula (before reading) for this phrase can be found in Diagram V:

$$+D_{spec}\pm Sb_{emp}\pm Sb_{quant}\pm Sb_{qual}\pm Sb_{ag}+H_{gl}\pm R_{nex}\pm R_{rel}\pm A_{rel}$$

It is possible to read every tagmeme in this syntagmeme, which might yield a phrase like,

> The only two major books of mine for you to read which are on the table, the blue ones,....

However, we read

(3) Phr_2: $D_{spec}Sb_{quant}Sb_{ag}H_{gl}$

In the grammar presented in this thesis we have only got as far as (3). Projecting further, we come to optional conjoining rules. As at clause level, these rules permit us to conjoin only K-equivalent tagmemes: that is, tagmemes whose grammatical form and grammatical meaning are the same. Thus, for example, Sb_{ag} can only be conjoined with Sb_{ag}. The grammar will not permit conjoinings which might yield strings like

> *These and his two books are lost.
>
> *Only and his two books are lost.

Without the complex tagmeme, including both form and meaning (category), these incorrect conjoinings would be generated. As we stated in Chapter III, constraints on conjoining are the major criterion for establishing categories of grammatical meaning. Further analysis would show that this criterion holds for all of the categories established for phrase level reading rules.

However, in the particular path we are following, phrase level conjoining options were not taken. In the phrase level permutation rules the tagmeme Sb_{ag} is obligatorily permuted to a position before Sb_{quant} or after H_{gl}, to prevent

*two his books....

Hence,

(4) $D_{spec}Sb_{ag}Sb_{quant}H_{gl}$ (or $D_{spec}Sb_{quant}H_{gl}Sb_{ag}$)

At this point we have generated the string,

$S_{gl}(D_{spec}Sb_{ag}Sb_{quant}H_{gl})Pr_{st}(....)C_{qual}(....)$

The ellipses in parentheses merely show that we did not provide exponents and readings for the predicate and complement. At this point the phrase level tagmemes in (4) are given exponents. The exponents for D_{spec} are the, a, and Ø. The selection of these exponents is context sensitive and will contain rules like

D_{spec} in context $(....Sb_{quant})$ = the/Ø

D_{spec} in context $(....Sb_{ag}H_{gl})$ = Ø

Let us suppose that the exponent for Sb_{ag} is W_1+poss, for Sb_{quant} is W_2, and for H_{gl} is W_3+pl in context $(Sb_{ag}Sb_{quant}...)$. In word level reading rules W_1 might yield Name, Phr_1, or Pronoun, plus poss. That is, the exponent of Sb_{ag} might be

John's

The king of England's

His

If a phrasal exponent is chosen, it is an instance of word level embedding and we must cycle back to the phrase level reading rules. In our example the selection is Prn+poss. At this point, note that the exponents are not necessarily tagmemes. Labels for word classes will be generated at the end of the grammatical process.

Let us assume that the reading of W_2 is Number and that for W_3+pl is $\mathrm{Prestem}_{act}\mathrm{Stem}_{gl}$+pl. Because the latter are still tagmemes, not labels for morpheme classes, we must go through further reading and exponence cycles until $\mathrm{Prestem}_{act}$ yields Verb and Stem_{gl} yields Noun, a selection from a number of possible alternatives. At this point the terminal grammatical string takes the following shape:

$$(5)\ S_{gl}(D_{spec}[0]Sb_{ag}[Prn+poss]Sb_{quant}[Number]H_{gl}$$
$$[Prestem_{act}(Verb)Stem_{gl}(Noun)+pl])Pr_{st}(....)C_{qual}(....)$$

It is this total string which will be provided to the lexicon and the phonology to yield, possibly,

His two notebooks are lost.

The string at (5), then, includes the sort of information that this grammar will provide for a lexicon. For example, let us examine the information that this grammar provides about Noun in the strong above. We are assuming, of course, that our grammar will break down compounds like 'notebook' into its constituents and that any special meaning of the compound which is not derivable from its constituent morphemes can be handled by a cross-reference in the lexicon. The symbol Noun above carries the following semantic constraints derived from its generative history: +gl in context +quant [number] +ag [poss] +spec plus all the context which will be provided by Pr_{st} and C_{qual} when they are fully realized.

Chomsky (1965, p. 85) provides the following markers in the syntactic component of his grammar for the subject in the string underlying

Sincerity may frighten the boy.

Sincerity [+N, - Count, + Abstract]

For the object, 'boy,' the grammar will provide the following markers:

boy [+N, +Count, +Animate, +Human]

These markers are selections from the features

Abstract - Concrete
Animate - Inanimate
Mass - Count
Common - Proper
Human - Nonhuman

They may be carried as either + or - features, as in the examples above. In our final exponence rule (5), the category label gl carries the equivalent of + Concrete. (As we have seen earlier (4.4.2.3), abstract nouns will be from the categories prop, act, mo, st, and qual.) The category label and exponent quant [number] is equivalent to +Count. The category label and exponent ag [poss] and +spec are equivalent to +Common. The symbol Noun in our string is neutral in regard to the features animate-inanimate (and consequently also to the features human-nonhuman); that is, the 'concrete,' 'common,' 'count' noun selected as a lexical realization of the string (5) may be animate or inanimate.

This information (and that provided by the remainder of the clause) will define a lexical distribution class. That is, all words in the lexicon carrying the features +gl +quant [number] +ag [poss] +spec (and others not here included) will share the potentiality of being a realization of Noun in the string (5) above.

4.7 Summary

In Chapter I we examined several previous definitions of the notion subject, particularly subject in English. As we said at the beginning of that chapter, our problem here has been to explain this notion subject in a tagmemic grammar and to describe the realization of this theoretical construct in a partial grammar of English. Our definition has been operational; that is, an English subject is any reading of the clause level grammatical form designated S in Matrix 1 (Diagram IV). A reading results in a tagmeme, e.g. S_{ag}, which is an intersection of grammatical form and meaning. The internal structure of this tagmeme is given in exponence rules which specify the filler forms of this tagmeme. A full realization of a tagmeme includes its grammatical form and meaning plus its lexical form and meaning (though we have only very crudely marked lexical meaning here): S^{x}_{ag}: Phr_1, Phr_2, Phr_3, Phr_4, RCl_1, RCl_2. Here the grammatical form is S (subject), the grammatical meaning is ag (agent). The lexical form will eventually be realized through successive reading and exponence operations symbolized here by the labels for filler types (e.g. Phr_1, Phr_2, etc.). Only at word level will actual labels for lexical classes appear, defined in terms of all preceding grammatical operations in a particular derivation. Lexical meaning has only been indicated by the superscript x marking only in some general way that two or more tagmemes must be realized lexically by items about which it is true that x is a y or x is identical with y. However, as we have remarked, the major function of the superscripts here is to keep before us the fact that there is an important omission in the grammar.

A crucial distinction in this grammar is the distinction between levels: a tagmeme is a grammatical construct operative only at a single level, each level being distinctive in the grammatical forms (relations) that it defines. Here we follow Pike, Longacre, and Lamb[13] in insisting on the grammatical relevance of a hierarchy of grammatical levels or strata.

As an operational definition of subject in English, the partial grammar above does not provide axiomatic procedures for the segmentation of actual clauses into their constituents. That is, we know of no formal way to operate the grammar backwards, so that its input is a particular clause and its output a grammatical analysis. One can, however, guess at the structure of a particular clause and then attempt to generate it, thus using the grammar heuristically; as people are much cleverer than grammars, this is not very difficult and reflects a human competence no formal grammar has yet explained.

In this partial grammar we hope also to have indicated (though not completely demonstrated) the generative capacity of a tagmemic grammar, most of which has been only implicit until recently. We have shown how such a grammar may be recursive, insisting on a distinction between linear recursiveness (conjoining) and nesting recursiveness (embedding) which has not been made explicitly before, as it assumes a generative grammar which clearly differentiates a hierarchy of levels.

We have shown the importance of differentiating also between surface and deep structure in a tagmemic grammar by formalizing the previously implicit difference between grammatical form (surface structure) and grammatical meaning (deep structure) in the two levels (clause and phrase) examined here.

By presenting a tagmemic grammar in a partially generative format (many crucial assumptions have not yet been formalized), we have made many of our claims testable. If there are subjects of English clauses which are not generated by the rules described above, then our partial grammar is incomplete if the description of these subjects does not contradict our rules, wrong if it does.

Writing a generative grammar, even a partial one, quickly reveals to the analyst that each part of the grammar depends on the whole. Hence, many of our rules for the English subject must be taken as speculative until a more complete grammar in this format is produced. This is particularly true of the predicate tagmemes we have posited.

Writing such a grammar, rather than a looser, more traditional one, also reveals clearly to the analyst what he does not know, for each step in the derivation must be explicit. We are quite aware that a great deal remains to be explained about the English subject. As but one example of this great residue, the following sentence has a subject which this analyst does not yet know how to describe:

> Not John's temerity but rather his willingness to reveal himself so often and with such unfortunate results is what embarrasses me.

Among other things, the subject here (S_{qual}) includes as part of its internal structure negation, pronominalization, conjoining with 'but,' and intensification. All of these clearly relevant grammatical features of the subject await further study.

FOOTNOTES

CHAPTER I

1. The relevance of this aspect of grammatical units is most clearly seen in Pike (1964) where situational roles and grammatical roles of tagmemes are distinguished as separate features. This article has been seminal for much of the theoretical material in this study.

2. This information is in many ways equivalent to that provided by Fillmore's cases. See Fillmore, 1967.

3. I can find no reference to these distinctions in Chomsky, though other transformational grammarians (e.g. Lees, 1964a,p.13) introduce constituents like Loc, Tm, Man into phrase structure rules, thus mixing, as does Longacre (1964a), symbols for grammatical relations with those for grammatical meaning. This results in no distinction between the relation "subject" and the grammatical meaning location in a sentence like

Here is where it is.

We will insist that these aspects remain distinct. (For further discussion, see section 2.3 below.)

4. The notion of equivalence chains is taken from Harris (1963). The notion of the domain of a chain is first developed, I believe, in Lockwood (1965).

5. Two formal computer analyses of anaphoric equivalence are those by John C. Olney (1966) and S.N. Jacobson (1966).

6. Harold King has pointed out that these four aspects (grammatical form, grammatical meaning, lexical form, and lexical meaning) are parallel to distinctions Bloomfield made in _Language_, p. 264:

> Meaningful unit of linguistic signalling, smallest or complex: _linguistic form_; the meaning of a linguistic form is a _linguistic meaning_;
> (a) lexical: _lexical form_; the meaning of a lexical form is a _lexical meaning_.
> (b) grammatical: _grammatical form_; the meaning of a grammatical form is a _grammatical meaning_.

7. This is one of Postal's repeated criticisms of tagmemics. See Postal (1964, pp. 35-37, and 1966, p. 94). Postal sees tagmemics as simply adding extra nodes to ordinary phrase structure trees. For a detailed response to Postal's main criticisms, see section 3.2.8 below. In part Postal's misunderstanding of tagmemics is caused by his not understanding (or accepting) the relevance of grammatical levels.

8. We are here distinguishing Longacre's S.C. analysis from Harris's (1964) "string analysis of sentence structure," which we interpret as a type of center-adjunct analysis, chiefly on the grounds that Harris does not assume grammatical levels in his analysis.

9. Note that this affinity is explained in S.C. analysis by relationships on two different levels. The predicate-object relationship of _eat_ + _your bread_ is a clause level relationship, while the possessive-

head relationship of _your_ + _bread_ is a phrase level relationship. This distinction is discussed in detail in section 4.1 below.

10. For a discussion of the differences between Harris' approach and Longacre's, see Longacre (1963).

11. Harris never explains what the properties of occurence of a sentence are. Clear indication that whole sentences (considered as units) do have such grammatically relevant properties (an assumption we share with Harris) can be found in Gunter (1966).

12. Note that we are talking about only elementary surface structure, not deep or underlying structure. Following Harris's system of analysis, the elementary structure of the following sentences is identical:

(a) He is slow in walking.
(b) He is slow in the morning.

13. This limited theoretical framework (or set of questions) is not meant as a sufficient evaluation procedure for a grammar. The focus here in Chapter I is on analysis rather than formal presentation; hence many relevant criteria of evaluation(e.g. those proposed by Chomsky, 1957) are not considered until later. What is sought here is a reasonable sytem for classifying and comparing a number of grammarians in reference to a single problem: the description of the English subject.

14. This technique of establishing lexical classes by substitution in a frame is discussed by Pike (1967, pp. 135,250) and Garvin (1964, p. 60ff.) and illustrated in detail in Hornby (1954) and Alexander and Kunz (1964). It is apparently rejected by Lees (1963, p.551), who writes,

> No matter how class membership is designated, whether honestly by listing or subscripting, or whether by the misleading technique known as substitution in frames....

Though I shall argue below (3.2.2) against the use of lexical substitution classes as a means of describing syntactic patterns, I can see no reason to call classes derived by substitution "misleading." Substitution continues to be one of our most useful discovery procedures.

15. Notice that the complexity of the following formulation is greatly reduced in most criticisms of "Freisian structuralism." What is criticized is often Paul Roberts' (1958) interpretation and simplification of Fries rather than Fries himself. Similar criticisms of Chomsky seem often to be based unfairly on Roberts' (1964) interpretation and simplification of transformational grammar.

16. For a discussion of the use of meaning (in several different senses) in grammatical analysis, see Pike (1967, pp. 276-280) This discussion ends with the statement, subscribed to here, that

> ...the linguist's task is not done when he knows -- or shows -- that we are saying something (in "well-formed" sentences) but only after he knows -- and shows -- approximately what we are saying. Thus our grammatical formulas must have ref-

> erence to grammatical meaning if we wish to generate sentences which are meaningful and usable by the speaker. If we generate the meaning of sentences, none of which we can use or join into larger relevant and structured sequences, why bother?

See also Pike (1967, pp. 149, 500) For an argument against the possible charge of circularity in an approach which assumes certain kinds of meaning in syntactic description, see Chomsky (1957, p. 56ff.).

17. This analysis in which and is seen as a constituent of one of the conjoined elements is surely wrong. For an alternative analysis, see 2.2 below.

18. The various kinds of rules used in transformational grammars are clearly explained in Fillmore (1965).

19. See Postal (1966) and Chomsky (1965).

20. Surely this is a very inconsistent listing, including main-verb as a syntactic relation of the same sort as subject.Much of the difficulty rests on the strong binary bias that transformational grammarians seem to inherit from traditional I.C. analysis, causing them to assume that the division between subject and predicate is on a different rank level from that between verb and object. There is very little motivation for this rank difference and few other grammarians put object at a different rank from subject. Note that Jespersen (1933, pp. 78-79) put the verb at lower rank, calling the subject and object both primaries; Tesniere (1966,pp.14-15) saw the subject and object as subordinate (secondary) to the verb. Both, however, put subject and object in the same rank level. For further discussion of rank

and nexus, see section 3.2.3 below.

21. Major categories are so identified in Lees (1962, p.12).

22. Chomsky writes (1965, p.95) that his new rules will " analyze a category into a complex symbol in terms of the frame in which this category appears." A similar approach is adopted in section 4.6 below. Harris (1962, p. 27) also adopts this approach to subclassification.

23. Chomsky writes (1965, p. 88): "It is no longer true that a phrase-marker can be represented as a labelled tree-diagram."

24. This seems at least a partial support for Pike's statement (1967, p. 505):

> At some point, it would seem, transform grammar needs a further unit to be identified and transformed. It is our belief that the tagmeme, or something very much like it in structure, is needed for this purpose in transform grammar.

With the additions suggested by Fillmore, transform grammar comes to resemble tagmemics even more, so that Fillmore can write (1967, p.109):

> There is an easy conversion from case-grammar diagrams to 'tagmemic' formulas too, as long as the case categories unarily dominate NPs (or S). Or, for that matter, a case-grammar diagram can simply be 'read off' as a tagmemic formula, as long as certain symbols are designated as function indicators. One can easily say "NP filling an A slot" as anything else. The crucial difference between the modification of transformational grammar that I have been suggesting and the typical tagmemic study is the insistence on discovering the 'deepest' level of 'deep structure.'

In section 2.3 below it will be argued, in effect, that tagmemic grammars can attain this 'deepest' level.

CHAPTER II

1. Most tagmemic grammars have been written so as to be productive. As Pike has stated (1957, p.121), tagmemic grammars make the assumption that an adequate analysis "should be fruitful for the productive manipulation of language." See also Pike (1967,pp. 280-281).

2. One feature of Longacre's model is left out here: he includes superscripts on tagmemes to mark that a tagmeme may be repeated within a particular reading. These superscripts will be discussed below in section 2.2.

3. See Postal (1964, pp. 33-51, and 1966) for a transformational perspective on tagmemics. Postal is partially answered in 3.2 below.

4. For a sample of Tamil grammar presented according to Longacre's model, see Zvelebil (1965).

5. This example is from Koutsoudas (1966, p. 249).

6. For a detailed discussion of the relation between but and negation, see Gleitman (1965, p. 266 ff.) and Longacre (1967).

7. Sentence level tagmemes are discussed in Longacre (1967). Levels beyond the sentence are discussed in Pike (1966, pp. 79-92) and Becker (1965, 1966a, 1966b).

8. Ross, Fillmore, Lakoff, and others have recently argued that "the parts of speech which have traditionally been called verbs and adjectives should really be looked upon as two subcategories of one major lexical category, predicate." (Ross, 1966b,p.1) Their arguement is very convincing but of little major importance here, for, as will be discussed in 3.2.2, lexical distribution classes (i.e. parts of speech) play little part in the grammar adopted here. See Fillmore (1967, p.50) and Lakoff (1965).

9. Sentence syntagmemes and tagmemes (as distinct from clause syntagmemes and tagmemes) are identified in Longacre (1967). See also Halliday (1967, p.80).

CHAPTER III

1. The universality of categories like agent, instrument, location, etc. is clearly and deeply argued in Fillmore (1967). Charles C. Fries has suggested (privately) that what we have called grammatical meaning (Aspect B) may well have close relation to Old English cases.

2. Besides the traditional distinction between notional or logical subjects (our aspect B) and grammatical subjects (our aspect A), there are several other meanings for the term subject, making it at times rather ambiguous. (1) Subject is the name for a grammatical relation or functional slot in a clause. (2) a subject tagmeme is more than just a relation but rather a complex grammatical unit with four aspects: grammatical form, grammatical meaning, lexical form, lexical meaning. Subject tagmemes include subject as agent, subject as proposition, subject as goal, subject as instrument, etc., plus the exponents or realizations of each. Often the grammatical form (subject) gives the name to the whole tagmeme; we might just as well let the grammatical meaning provide the name and speak of agent as subject, agent as complement, agent as adjunct -- all agent tagmemes. Such a focus is useful in discourse analysis where we wish to follow a particular category of grammatical meaning through a number of sentences. (3) When we speak of the English subject or subject in English, we are speaking of a class of tagmemes, i.e. all the tagmemes whose grammatical form is subject. We have tried to keep these different meanings clear in our discussion.

3. Words like fact, result, idea, plan that we label category words for the category prop are called sentence-names (N_s) by Harris (1964, p.29). This subject tagmeme (S_{prop}) is similar to his object type 8 (p.31). This pattern is also Hornby's Noun Pattern 3 (1954, pp. 132-134).

4. Our category goal is very similar to Fillmore's objective case (1967, p.47). Fillmore states that this case is

> the semantically most neutral case, the case of anything representable by a noun whose role in the action or state identified by the verb is identified by the semantic interpretation of the verb itself; conceivably the concept should be limited to things which are affected by the action or state identified by the verb. The term is not to be confused with the notion Direct Object, nor with the name of the surface case synonymous with Accusative.

5. The categories nexus and proposition both represent what might be called semantic embedding (in contrast to grammatical embeddding). That is, categories represent linguistic categorization of the speaker's (or hearer's) experience, situational roles in Pike's terms (1964). In these two categories language itself becomes part of experience or the situation. See also Jakobson (1957) where he discusses instances of "the code referring to the message" (p.2).

6. Our three predicate categories (act, mo, st)overlap considerably with Halliday's (1967,p.39) directed act, non-directed act, and ascription, which he calls "process types."

7. Lees (1965) discusses this problem of the definition of sets of objects for which there is no formal test for membership. We would argue, however, that though there may be no completely rigorous test for membership in such a set, there are criteria which, when used intelligently, strongly indicate membership in a set and which are extremely useful to an analyst in overcoming his own inevitable biases. It is these heuristic criteria that will be discussed below.

8. The similarity of our three categories (*act*, *mo*, *st*)to Halliday's process types (see footnote 6 above) provides some outside support for our analysis. Fillmore (1967), on the other hand, does not include the verb within the case system but rather uses the cases of noun phrases to select the verb. Chomsky (1965) and Harris (1964) take a similar approach. This approach seems to seriously confuse deep and surface structure; that is, Fillmore's deep structure strings (1967, p.45) have the form

$$\text{Verb} + \text{Case}_1 + \ldots + C_n$$

We would argue that "Verb" (our Pr) is as much a part of the surface structure as subject, object, or complement.

9. Tagmemic studies of discourse analysis are presently in progress. These studies grow, in large part, out of Pike (1964). Preliminary studies of paragraph structure may be found in Becker (1965 and 1966b). Relevant non-tagmemic work includes Ross (1966a), Jacobson (1966), and Olney (1966).

10. Nida (1960, pp. 13-14) states:

> Jespersen is really the first to approach the problems of larger patterns in a serious way. The difficulty in his approach is (1) the manner of classification which is based on a semantic rather than a formal relationship between constituents, and (2) the interpretation of the difference between morphology and syntax as an inner and an outer approach. The making of an attributive _barking_ in _the barking dog_ of the same rank as _barks_ in _the dog barks_ because of the parallelism in meaning value does not account for the difference in endocentric and exocentric constructions which Bloomfield so adequately demonstrates.

Lees (1960, pp. xxiii - xxv) also rejects Jespersen's analysis because

> he did not conceive of grammatical relations as autonomous, formal features of sentences which could be studied independently.

11. Pike (1966, pp. 13-14, 57-60) shows similar obligatory exclusion of tagmemes in some African languages.

12. This statement should be qualified. The distinction between surface structure and deep structure refers both (1) to a distinction between representations of sentence structure within a particular theory, transformational grammar, and (2) to a more general distinction between an underlying "structural meaning" on the one hand and various surface representations of this meaning on the other, a distinction found, for instance, in frequent discussions of the difference between "logical" subject and "grammatical" subject. In its dual notation (e.g. subject as actor) tagmemics has preserved this distinction, as it has in the continual insistence that grammatical units be considered form-meaning composites. Thus the distinction between deep and surface structure is strongly motivated in tagmemic theory (see Pike, 1967, pp.219, 276-280). However, as

was argued in section 2.2, this distinction has been blurred in some tagmemic grammars. (That it has also been blurred in transformational grammars is suggested in footnote 8 above.)

CHAPTER IV

1. It has been suggested that the following sentence includes an instance of H_{st}:

His existence was threatened.

Though this may turn out to be the correct analysis, we advocate caution here. It is important to distinguish between paraphrase and equivalence of grammatical meaning. The following two sentences can be considered roughly paraphrases of each other, though they are not equivalent in grammatical meaning:

John saw the book. ($S_{ag}Pr_{act}O_{gl}$)

John looked at the book. ($S_{ag}Pr_{act}Ad_{loc}$)

Though exist and be may occur as paraphrases of each other, they likewise occur in distinctly different syntagmemes. For example, sentences with be always require complements, though this is not the case with exist:

Dinosaurs existed.

*Dinosaurs were.

Hence we would say that Pr_{st} is always followed by C. Similarly, we have seen (section 3.2.9) that Pr_{st} may occur with any subject tagmeme. This is not the case if exist is an exponent of Pr_{st}, for S_{nex} does not occur before exist:

For John to go would be a shame.

*For John to go would exist as a shame.

Any solution of this problem awaits further analysis of the predicate, however.

2. It has been suggested that the following sentence includes an instance of C_{quant}:

The reasons are many.

We would analyze this sentence as including C_{qual} on the basis of conjoining potential, one of our major criteria for establishing contrasting categories (section 3.2.1):

The reasons are many and interesting.

See footnote 1 above for the distinction between paraphrase and grammatical equivalence.

3. Matrix derivation is described and illustrated for Potawatomi and Arabic in Pike and Erickson (1964), for Navajo in Pike and Becker (1964), and for English in Becker (1966a).

4. For a strong objection to this analysis, see Longacre (1963, p.477). If we adopt this analysis, a passive syntagmeme would now be read

$$S_{gl}\ Pr_{st}\ C_{act}\ \pm Ad_{ag}\ldots$$

This is similar to Harris's analysis of the passive (1964, p.33).

5. The derivation of many pre-nominal modifiers from post-verbal adjectives in relative clauses is, of course, common in transformational grammars. The process is clearly described in Lees (1962).

6. This is further support for the interpretation of there as a dummy subject. See section 3.2.6 above.

7.

8. See Pike (1945, pp. 30-33).

9. See Lockwood (1965), Ross (1966a), and Becker (1965 and 1966b) for discussion of the domains of anaphoric relations.

10. Fillmore (1967, p.60 ff) interprets all noun phrases as marked by prepositions which are deleted in certain surface relations. These prepositions function as case markers. Here we interpret prepositional phrases (relative phrases) as derived from phrase level syntagmemes.

11. Numbers (5) and (6) in the diagram are allo-syntagmemes of Phr_5. Their only difference is that in (5) the appositive is a realization of prop while in (6) the appositive is a realization of nex. This difference is discussed in section 4.4.3. This single structural difference is not sufficient to contrast them as separate syntagmemes. Numbers (4) and (10), however, are different syntagmemes even though both have as a nucleus H_{qual}. They are separated on the basis of internal structure (the presence of man and ag) and external distribution (4 is an exponent of S_{ag} and S_{gl}, 10 is an exponent of S_{qual}). Criteria for emic contrast are discussed in section 3.2.1.

12. It should be noted that the generative process has begun somewhere in the middle in this grammar. Clause syntagmeme (our initial units) are realizations of sentence level tagmemes (described in Longacre, 1967) and sentences themselves are realizations of higher level tagmemes up to the level of discourse. These higher level patterns are discussed in Pike (1966, pp. 79-92) and Becker (1965, 1966a,and 1966b).

13. For arguments on the relevance of grammatical levels in linguistic description, see Longacre (1960, 1964a, 1964b), Pike (1967, pp.437-443, 479-481), and Lamb (1965).

BIBLIOGRAPHY

D. Alexander and W. J. Kunz (1964) *Some Classes of Verbs* University of Indiana Linguistics Research Project. (Principal investigator F. W. Householder, Jr., assistant director, P. H. Matthews.)

Leonard Bloomfield (1933) *Language*. New York: Holt, Rinehart and Winston.

A. L. Becker (1965) "A Tagmemic Approach to Paragraph Analysis" *College Composition and Communication* 16. 237-242.

A. L. Becker (1966a) "Item and Field: A Way into Complexity" in Carol J. Kreidler, ed., *On Teaching English to Speakers of Other Languages* (Series II) Champaign: National Council of Teachers of English. 132-138.

A. L. Becker (1966b) "Symposium o Paragraphs" *College Composition and Communication* 17. 67-72.

A. L. Becker (1967) "Conjoining in a Tagmemic Grammar of English" *Monograph Series on Languages and Linguistics* 20 Washington: Georgetown University (In press.)

A. L. Becker and Richard E. Young (1965) "Toward a Modern Theory of Rhetoric: A Tagmemic Contribution" *Harvard Educational Review* 35. 4. 450-468.

Noam Chomsky (1957) *Syntactic Structures* (Janua Linguarum, No. 4) 'S-Gravenhage: Mouton.

Noam Chomsky (1962) "A Transformational Approach to Syntax" in Fodor and Katz (1964) 211-245.

Noam Chomsky (1965) *Aspects of the Theory of Syntax* Cambridge: The M.I.T. Press.

Paul Christophersen (1939) *The Articles, A Study of their Theory and Use in English* Copenhagen: Einar Munksgaard.

Ruth H. Crymes (1965) Some *Systems of Substitution Correlation in Modern American English with Implications for the Teaching of English as a Second Language* (Columbia Ph.D. Dissertation).

Charles J. Fillmore (1965) *Indirect Object Constructions in English and the Ordering of Transformations* (Monographs on Linguistic Analysis No. 1) The Hague: Mouton.

Charles J. Fillmore (1966) "A Proposal Concerning English Prepositions" *Report of the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies* (Georgetown Monograph Series on Languages and Linguistics) 19. 19-33.

Charles J. Fillmore (1967) "The Case for Case" To appear in Emmon Bach and Robert Harms, eds., *Proceedings of the Texas Symposium on Language Universals April 13-15, 1967* (New York: Holt, Rinehart and Winston)

Jerry A. Fodor and Jerrold J. Katz (1964) *The Structure of Language, Readings in the Philosophy of Language* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Charles C. Fries (1940) *American English Grammar*,New York: Appleton-Century.

Charles C. Fries (1952) *The Structure of English* New York: Harcourt, Brace.

Paul L. Garvin (1964) "Syntactic Units and Operations" in On Linguistic Method (Janua Linguarum No. 30) The Hague: Mouton.

Paul L. Garvin and Bernard Spolsky, eds. (1966) Computation in Linguistics: A Case Book Bloomington: Indiana University Press.

Lila R. Gleitman (1965) "Coordinating Conjunctions in English" Language 41. 260-293.

Richard Gunter (1966) "On the Placement of Accent in Dialogue: A Feature of Context Grammar" Journal of Linguistics 2.2. 159-180.

Barbara C. Hall (196[illegible]) Subject and Object in English (M.I.T. Ph.D. Dissertation)

M. A. K. Halliday (1967) "Notes on Transitivity and Theme in English, Part I" Journal of Linguistics 3.1. 37-81.

Zellig S. Harris (1962) String Analysis of Sentence Structure (Papers on Formal Linguistics, No. 1) The Hague: Mouton.

Zellig S. Harris (1963) Discourse Analysis Reprints (Papers on Formal Linguistics, No. 2) The Hague: Mouton.

Archibald A. Hill (1966) "A Re-examination of the English Articles" Report of the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies (Georgetown Monograph Series on Languages and Linguistics) 19. 217-231.

Louis Hjelmslev (1959) "La Nature du Pronom" in Essais Linguistiques (Travaux du Cercle Linguistique de Copenhague) XIII.

A. S. Hornby (1954) A Guide to Patterns and Usage in English London: Oxford University Press.

Roman Jakobson (1939) "Signe Zéro" Mélanges de Linguistique Objects à Charles Bally Genève 143-52.

Roman Jakobson (1957) "Shifters, Verbal Categories, and the Russian Verb" Cambridge: Harvard Department of Slavic Language and Literature.

S. N. Jacobson (1966) "A Modifiable Routine for Connecting Related Sentences of English Text" Chapter 11 of Garvin and Spolsky (1966).

Otto Jespersen (1923) Language; Its Nature, Origin, and Development New York: Henry Holt.

Otto Jespersen (1927) A Modern English Grammar, Part III, Syntax Second Volume. Heidelberg: Carl Winters.

Otto Jespersen (1933) Essentials of English Grammar New York: Henry Holt.

Otto Jespersen (1934) The Philosophy of Grammar New York: Henry Holt.

Otto Jespersen (1937) Analytic Syntax Copenhagen: Munksgaard.

Rolf Karlsen (1959) Studies in the Connection of Clauses in Current English Bergen: J. W. Eides Boktrykkeri A. S.

Edward S. Klima (1959) "Negation in English" in Fodor and Katz (1964) 246-323.

Andreas Koutsoudas (1966) Writing Transformational Grammars: an Introduction New York: McGraw-Hill.

George Lakoff (1965) "On the Nature of Syntactic Irregularity" Harvard Computation Laboratory Report No. NSF-16.

Sydney M. Lamb (1964) "On Alternation, Transformation, Realization, and Stratification" Monograph Series on Languages and Linguistics 17. 105-122.

Robert B. Lees (1960) The Grammar of English Nominalizations IJAL (Publication 12 of the Indiana University Research Center in Anthropology, Folklore and Linguistics) 26.3 (Part II).

Robert B. Lees (1961) "The Constituent Structure of Noun Phrases" American Speech XXXVI. 159-168.

Robert B. Lees (1962) "The Grammatical Basis of Some Semantic Notions" Report of the Eleventh Annual Round Table Meeting on Linguistics and Language Study (Georgetown Monograph Series on Languages and Linguistics) 13. 5-20.

Robert B. Lees (1963) "Review of J. Nemeth, Turkish Grammar" Language 39. 548-56.

Robert B. Lees (1965) "On the Testability of Linguistic Predicates" Linguistics 12. 37-48.

Nguyen Dang Liem (1966) English Grammar: A Combined Tagmemic and Transformational Approach (Linguistic Circle of Canberra Publications, Series C, No. 3) Canberra: The Australian National University.

David Lockwood (1965) "Pronoun Concord Domains in English: An Inquiry into Grammatical Structure above the Sentence" (Paper delivered to Linguistic Society of America in Ann Arbor, Summer, 1965).

Robert E. Longacre (1960) "String Constituent Analysis" Language 36.1. 63-88.

Robert E. Longacre (1963) Review of Harris (1962) in Language 39.3 (part 1). 473-478.

Robert E. Longacre (1964a) Grammar Discovery Procedures: A Field Manual (Janua Linguarum, No. 33) The Hague: Mouton.

Robert E. Longacre (1964b) "Prolegomena to Lexical Structure" Linguistics 5. 5-24.

Robert E. Longacre (1965) "Some Fundamental Insights of Tagmemics." Language 41. 65-76.

Robert E. Longacre (1966) "Trique Clause and Sentence: A Study in Contrast, Variation and Distribution" IJAL 32.3. 242-252.

Robert E. Longacre (1967) "The Notion of Sentence" Monograph Series on Languages and Linguistics 20 Washington: Georgetown University (In press).

Eugene A. Nida (1960) A Synopsis of English Syntax Norman, Oklahoma: Summer Institute of Linguistics. (All except Preface is 1943 dissertation.)

J. C. Olney (1966) "A Research Plan for Investigating English Discourse Structure with Particular Attention to Anaphoric Relationships" (System Development Corporation, Technical Memo (L)-3256).

Kenneth L. Pike (1943) "Taxemes and Immediate Constituents" Language 19.2. 65-82.

Kenneth L. Pike (1945) The Intonation of American English Ann Arbor: The University of Michigan Press.

Kenneth L. Pike (1957) "Grammemic Theory in Restricted Problems of Morpheme Classes" IJAL 23.3. 119-128.

Kenneth L. Pike (1962) "Dimensions of Grammatical Constructions" Language 38. 221-44.

Kenneth L. Pike (1963) "Theoretical Implications of Matrix Permutation in Fore (New Guinea)" Anthropological Linguistics 5.8. 1-23.

Kenneth L. Pike (1964) "Discourse Structure and Tagmeme Matrices" Oceanic Lingusitics 3. 5-25.

Kenneth L. Pike (1966) Tagmemic and Matrix Linguistics Applied to Selected African Languages (Final Report for U.S. Office of Education Contract No. OE-5-14-065) Ann Arbor: Center for Research on Language and Language Behavior.

Kenneth L. Pike (1967) Language in Relation to a Unified Theory of the Structure of Human Behavior 2nd Edition (Janua Linguarum, No. 24) The Hague: Mouton. (Originally published in 1954, 1955, 1960).

Kenneth L. Pike and A. L. Becker (1964) "Progressive Neutralization in Dimensions of Navajo Stem Matrices" IJAL 30. 144-54.

Kenneth L. Pike and Barbara Erickson (1964) "Conflated Field Structures in Potawatomi and in Arabic" IJAL 30.3. 201-212.

Richard S. Pittman (1948) "Nuclear Structures in Linguistics" Language 24. 287-92.

Paul Postal (1964) Constituent Structure: A Study of Contemporary Models of Syntactic Description. IJAL (Publication 30 of the Indiana University Research Center in Anthropology, Folklore and Linguistics) 30.1 (Part III).

Paul Postal (1966a) Review of Longacre (1964a) in IJAL 32.1 (Part 1). 93-98.

Paul M. Postal (1966b) "Underlying and Superficial Linguistic Structure" *Language and Learning* (A revision and expansion of the 1964 Special Issue of the *Harvard Educational Review*) New York: Harcourt, Brace and World. 153-175.

Paul Roberts (1958) *Understanding English* New York: Harper.

Paul Roberts (1964) *English Syntax* New York: Harcourt, Brace and World.

John Robert Ross ,(1966a) "On the Cyclic Nature of English Pronominalization" to appear in *Festschrift for Roman Jakobson*, Mouton, 1967.

John Robert Ross (1966b) "Adjectives as Noun Phrases" mimeographed copy of paper presented at the Winter Meeting of the Linguistic Society of America in New York, December 28, 1966.

James Sledd (1959) *A Short Introduction to English Grammar*. Chicago: Scott, Foresman.

Carlota S. Smith (1964) "Determiners and Relative Clauses in a Generative Grammar of English" *Language* 40.1. 37-52.

Lue~~i~~en Tesniere (1966) *Elements de Syntaxe Structurale* Paris: Librairie C. Klincksieck.

Rulon S. Wells (1947) "Immediate Constituents" *Language* 23. 81-117.

Benjamin L. Worf (1945) "Grammatical Categories" *Language* xxl. 1-5.

Kamil Zuelebil (1965) "Towards a Taxonomic and Generative Grammar of Tamil" *Archiv Orientalni* 33. 602-613.